

ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΥΠΟΛΟΓΙΣΤΩΝ



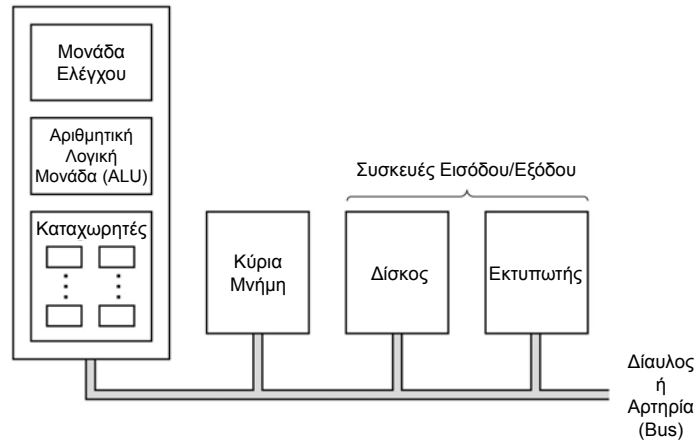
Διάρθρωση

1. Οργάνωση της μνήμης
2. Κύρια μνήμη
3. Κρυφή μνήμη
4. Κώδικες διόρθωσης λαθών
5. Δευτερεύουσα μνήμη
6. Μονάδες εισόδου/εξόδου



Οργάνωση Υπολογιστή

Κεντρική Μονάδα Επεξεργασίας
Central Processing Unit (CPU)



Τυπική οργάνωση υπολογιστή βασισμένου σε δίαυλο

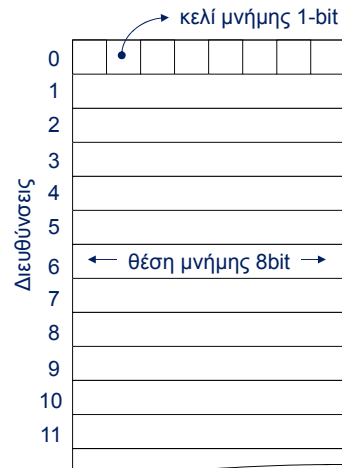


ΟΡΓΑΝΩΣΗ ΤΗΣ ΜΝΗΜΗΣ

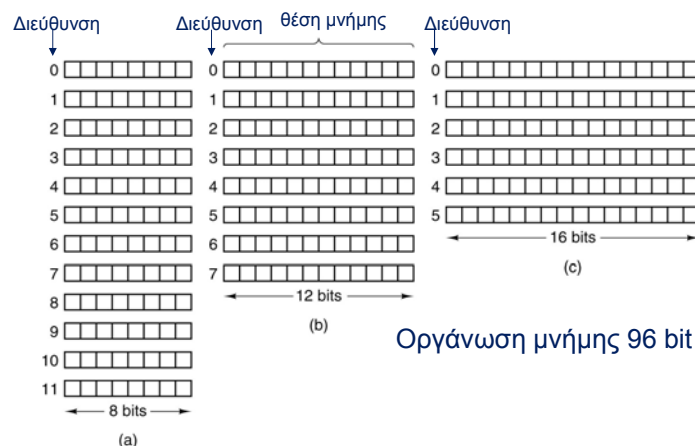


Λογική Οργάνωση Μνήμης

- Η μνήμη έχει τη λογική οργάνωση ενός πίνακα ο οποίος αποτελείται από κελιά.
- Συνήθως σε κάθε κελί αποθηκεύεται ένα bit πληροφορίας.
- Τα κελιά οργανώνονται σε θέσεις μνήμης όπου κάθε μία προσδιορίζεται με αποκλειστικό τρόπο από μία διεύθυνση (address).
- Με m bits για τη διεύθυνση μπορούμε να προσδιορίσουμε 2^m θέσεις μνήμης.



Οργάνωση της Μνήμης



- Λίγα bit ανά θέση $\Rightarrow$ πολλές θέσεις μνήμης και αντίστροφα.
- Μια θέση μνήμης είναι η μικρότερη προσπελάσιμη μονάδα.



Byte και Διάταξη στη Μνήμη

- Θέσεις μνήμης των 8 bit συνθέτουν ένα *byte*.

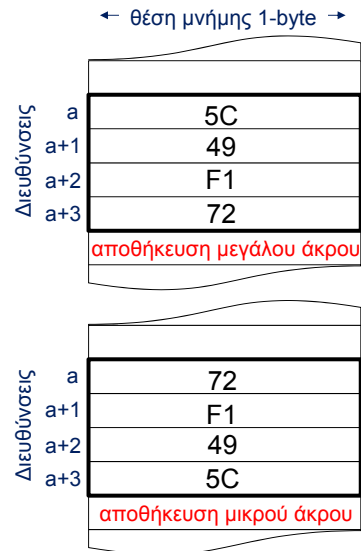
- Σήμερα οι επεξεργαστές διαχειρίζονται την πληροφορία σε *λέξεις (words)* των 32-bit ή 64-bit.

- Συνεπώς απαιτούνται πολλές θέσεις μνήμης του 1-byte για την αποθήκευση μιας λέξης, π.χ. για 32-bit θέλουμε 4 byte.

- μνήμη μεγάλου άκρου (big endian)
- μνήμη μικρού άκρου (little endian)

π.χ. αποθήκευση της 32-bit λέξης:

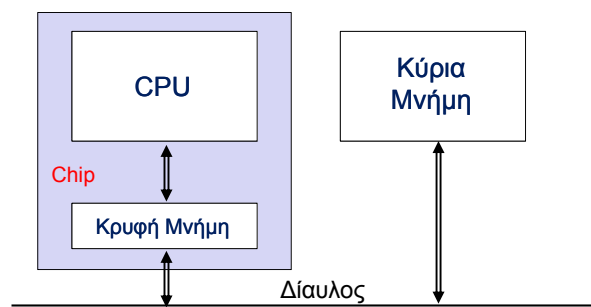
(5C 49 F1 72)_{Hex}
 MSByte LSByte



Οργάνωση Υπολογιστικών Συστημάτων – II

7

Κρυφή ή Λανθάνουσα Μνήμη



Χρήση μιας μικρής, γρήγορης αλλά υψηλού κόστους μνήμης μεταξύ της κύριας μνήμης και του επεξεργαστή, η οποία ονομάζεται *κρυφή μνήμη (cache)*.

Αποθηκεύει προσωρινά, πληροφορία συχνά χρησιμοποιούμενη από τον επεξεργαστή.



Οργάνωση Υπολογιστικών Συστημάτων – II

8

Αρχή της Τοπικότητας

- **Χωρική τοπικότητα (spatial locality)**: αν η CPU προσπελάσει μια θέση μνήμης, τότε με μεγάλη πιθανότητα θα προσπελάσει και τις γειτονικές της.
- **Χρονική τοπικότητα (temporal locality)**: αν η CPU προσπελάσει μια θέση μνήμης, τότε με μεγάλη πιθανότητα θα την ξαναπροσπελάσει άμεσα στο μέλλον.
- Αρχές λειτουργίας της cache:
 - Δεδομένα που τοποθετούνται στην cache διατηρούνται για μεγάλο χρονικό διάστημα (λόγω χρονικής τοπικότητας).
 - Στην cache μεταφέρονται ολόκληρες **ενότητες (block)** δεδομένων και όχι απλές λέξεις (λόγω χωρικής τοπικότητας).
 - ✓ Οι αποθηκευμένες ενότητες στην cache ονομάζονται **γραμμές κρυφής μνήμης (cache lines)**.
 - ✓ Π.χ. cache με γραμμή 64byte: μια αναφορά στη διεύθυνση μνήμης 260 θα επιφέρει μεταφορά των διευθύνσεων από 256 – 319.



Cache και Απόδοση

Έστω ότι ο επεξεργαστής προσπελαύνει μια λέξη k φορές σε σύντομο χρονικό διάστημα. Με τη χρήση cache, θα γίνει 1 προσπέλαση στην αργή κύρια μνήμη και $k-1$ προσπελάσεις στην γρήγορη cache μνήμη.

Το κέρδος αναλύεται ακολούθως, όπου ορίζουμε ως:

- t_c το χρόνο προσπέλασης της cache
- t_m το χρόνο προσπέλασης της κύριας μνήμης ($t_m \gg t_c$)
- h το **λόγο επιτυχιών (hit ratio)**, το ποσοστό των προσπελάσεων που μπορούν να ικανοποιηθούν από την cache. Ισχύει:

$$h = (k-1)/k$$

Αντίστοιχα ο **λόγος αποτυχιών (miss ratio)** είναι $1-h$.

Ο μέσος χρόνος προσπέλασης t_{av} που “βλέπει” η CPU είναι:

$$t_{av} = ht_c + (1-h)(t_c + t_m) = t_c + (1-h)t_m$$

Π.χ. με $t_c=20\text{ns}$, $t_m=100\text{ns}$ και $h=0.85$, ο μέσος χρόνος προσπέλασης είναι: $t_{av}=35\text{ns}$.



Θέματα Δομής και Οργάνωσης Cache

- Χωρητικότητα της κρυφής μνήμης, καθώς το κόστος είναι υψηλό.
- Μέγεθος της γραμμής κρυφής μνήμης, καθώς η καθυστέρηση μεταφοράς ενός block από την κύρια μνήμη είναι σημαντική.
- Οργάνωση της κρυφής μνήμης για τη διαχείριση της πληροφορίας.
- Τρόπος αποθήκευσης των δεδομένων στην κρυφή μνήμη:
 - *Ενοποιημένος (unified)*, εντολές & τελούμενα στο ίδιο τμήμα
 - *Διαμερισμένος (split)*, ξεχωριστά τμήματα για εντολές & τελούμενα
- Αριθμός κρυφών μνημών, καθώς χρησιμοποιούνται ευρύτατα σήμερα πολλαπλά επίπεδα κρυφής μνήμης, εντός και εκτός του τσιπ του επεξεργαστή.



Κώδικες Διόρθωσης Λαθών 1/3

- Για την προστασία της πληροφορίας χρησιμοποιούνται κώδικες ανίχνευσης ή διόρθωσης λαθών.
- Στα m bit δεδομένων προσθέτονται r bit πλεονασμού και σχηματίζεται μια *κωδικολέξη (codeword)* των n bit.
- Από τις 2^n δυνατές λέξεις μόνο 2^m είναι αποδεκτές. Αν κατά την ανάγνωση της μνήμης προκύψει μη αποδεκτή κωδικολέξη τότε έχει ανιχνευθεί λάθος στα δεδομένα της μνήμης.
- Ορίζουμε ως *απόσταση Hamming* μεταξύ δύο κωδικολέξεων τον αριθμό των θέσεων που διαφέρουν τα bit τους (βρίσκεται εύκολα με απλό XOR των bit κατά θέση και καταμέτρηση των 1 που προκύπτουν).
- Η ελάχιστη απόσταση Hamming μεταξύ των ζευγών κωδικολέξεων αποτελεί την *απόσταση Hamming του χρησιμοποιούμενου κώδικα*.

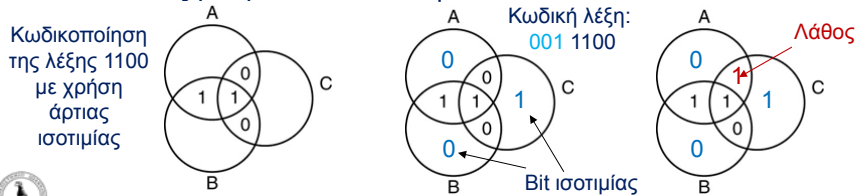


Richard Hamming
1915 – 1998



Κώδικες Διόρθωσης Λαθών 2/3

- Για την ανίχνευση d λαθών του ενός bit απαιτείται κώδικας με απόσταση Hamming $d+1$. Σε αυτή την περίπτωση, d λάθη του 1bit δεν μπορούν να μετατρέψουν μία κωδικολέξη σε άλλη κωδικολέξη.
 - Απλό παράδειγμα κώδικα για την ανίχνευση μεμονωμένων λαθών είναι η χρήση ενός επιπλέον *bit* **ισοτιμίας** (*parity bit*), το οποίο επιλέγεται έτσι ώστε το πλήθος των bit της κωδικολέξης με τιμή 1 να είναι άρτιο (ή περιττό).
- Για τη διόρθωση d λαθών του ενός bit απαιτείται κώδικας με απόσταση $2d+1$. Αυτή η συνθήκη εξασφαλίζει ότι αν συμβούν d λάθη η αρχική (σωστή) κωδικολέξη θα έχει μικρότερη απόσταση από την τελική μη κωδικολέξη σε σχέση με οποιαδήποτε άλλη κωδικολέξη και συνεπώς μπορεί να ταυτοποιηθεί!



Οργάνωση Υπολογιστικών Συστημάτων – II

13

Κώδικες Διόρθωσης Λαθών 3/3

Έστω κωδικολέξη των n bit, με m bit πληροφορίας και r bit πλεονασμού ($n=m+r$) για τη διόρθωση απλών λαθών του ενός bit.

- Κάθε μία από τις 2^m κωδικολέξεις έχει n μη αποδεκτές κωδικολέξεις με απόσταση Hamming 1 από αυτή (προσδιορίζονται με αντιστροφή ενός bit της κωδικολέξης κάθε φορά).
- Καμία άλλη κωδικολέξη δεν έχει απόσταση 1 από τις συγκεκριμένες μη αποδεκτές κωδικολέξεις.
- Καθώς ο συνολικός αριθμός λέξεων (κωδικολέξεων και μη) είναι 2^n θα πρέπει να ισχύει $(n+1)2^m \leq 2^n$. Με δεδομένο ότι $n=m+r$ η προηγούμενη ανισότητα ισοδυναμεί με $(m+r+1) \leq 2^r$.

m	r	n	Επιβάρυνση %
8	4	12	50
16	5	21	31
32	6	38	19
64	7	71	11
128	8	136	6
256	9	265	4
512	10	522	2

Bit πλεονασμού για τη διόρθωση μεμονωμένων λαθών

Οργάνωση Υπολογιστικών Συστημάτων – II

14

Αλγόριθμος Hamming 1/3

Σε έναν κώδικα Hamming προστίθενται r bit ισοτιμίας σε μια λέξη των m bit. Η νέα λέξη έχει $m+r$ bit. Εργαζόμαστε ως ακολούθως :

- Τα bit αριθμούνται από το αριστερότερο bit (MSB).
- Όλα τα bit που ο αριθμός τους είναι δύναμη του 2 είναι bit ισοτιμίας. Τα υπόλοιπα είναι bit δεδομένων.
- Ένα bit b_i της κωδικοποιημένης λέξης ελέγχεται ως προς την ισοτιμία του από τα bit ισοτιμίας $b_{2^0}, b_{2^1}, \dots, b_{2^j}$ ($2^j \leq i$), έτσι ώστε να ισχύει:

$$2^0 + 2^1 + \dots + 2^j = i \Leftrightarrow 1 + 2 + \dots + 2^j = i$$

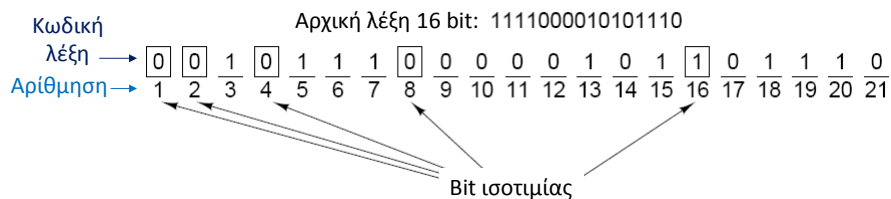
- Προφανώς, αν το bit b_i είναι bit ισοτιμίας τότε ελέγχεται από τον εαυτό του και μόνο.
- Η τιμή του κάθε bit ισοτιμίας καθορίζεται από τις τιμές των bit που ελέγχει σύμφωνα με τον τύπο της ισοτιμίας που ζητάμε (άρτια ή περιπτή).



Αλγόριθμος Hamming 2/3

Παράδειγμα κωδικοποίησης λέξης $m=16$ bit με $r=5$ bit ισοτιμίας για την ανίχνευση απλών λαθών (σύνολο $n=21$ bit):

- Τα bit 1, 2, 4, 8 και 16 είναι bit ισοτιμίας.
- Χρήση άρτιας ισοτιμίας (*αυθαίρετη επιλογή*).
- Το bit 1 ελέγχει τα bit: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21.
- Το bit 2 ελέγχει τα bit: 2, 3, 6, 7, 10, 11, 14, 15, 18, 19.
- Το bit 4 ελέγχει τα bit: 4, 5, 6, 7, 12, 13, 14, 15, 20, 21.
- Το bit 8 ελέγχει τα bit: 8, 9, 10, 11, 12, 13, 14, 15.
- Το bit 16 ελέγχει τα bit: 16, 17, 18, 19, 20, 21.

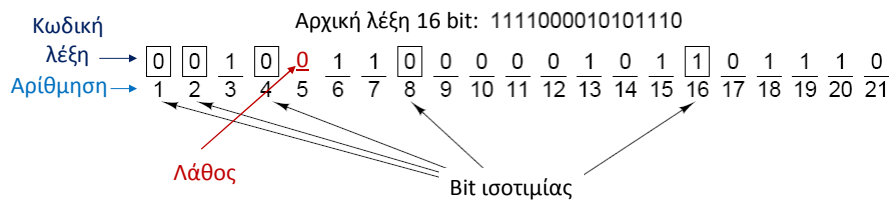


Αλγόριθμος Hamming 3/3

Αν εξαιτίας κάποιου σφάλματος το 5^ο bit είναι λανθασμένο (1→0) τότε:

- Το bit ισοτιμίας 1 είναι **λανθασμένο** (τα ελεγχόμενα bit περιέχουν πέντε 1).
- Το bit ισοτιμίας 2 είναι σωστό (τα ελεγχόμενα bit περιέχουν έξι 1).
- Το bit ισοτιμίας 4 είναι **λανθασμένο** (τα ελεγχόμενα bit περιέχουν πέντε 1).
- Το bit ισοτιμίας 8 είναι σωστό (τα ελεγχόμενα bit περιέχουν δύο 1).
- Το bit ισοτιμίας 16 είναι σωστό (τα ελεγχόμενα bit περιέχουν τέσσερις 1).

Το κοινό ελεγχόμενο bit των bit ισοτιμίας 1 και 4 είναι το 5^ο bit το οποίο και φέρει τη λανθασμένη τιμή. Άλλωστε $4+1=5$!



ΔΕΥΤΕΡΕΥΟΥΣΑ ΜΝΗΜΗ

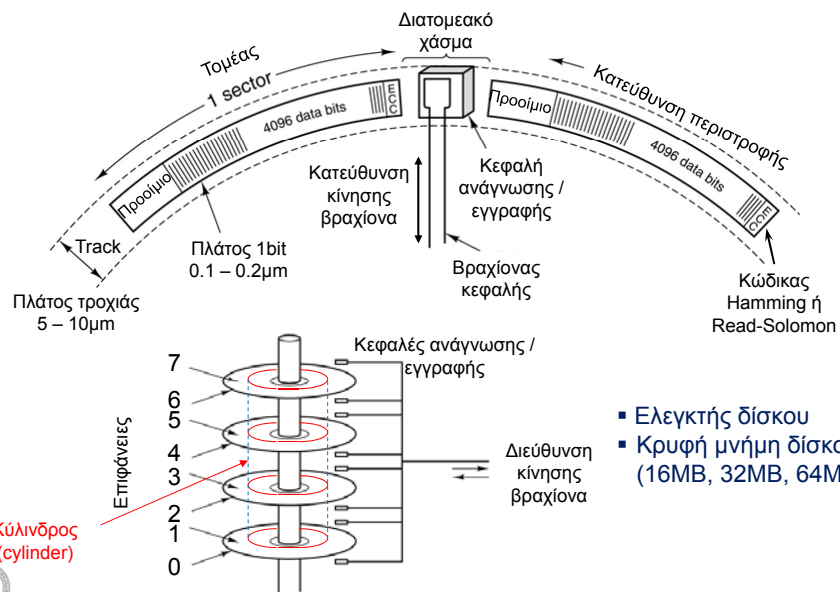


Ιεραρχία Μνήμης

Μείωση του κόστους ανά bit
Αύξηση της χωρητικότητας
Αύξηση του χρόνου προσπέλασης



Μαγνητικοί Δίσκοι



- Ελεγκτής δίσκου
- Κρυφή μνήμη δίσκου (16MB, 32MB, 64MB)



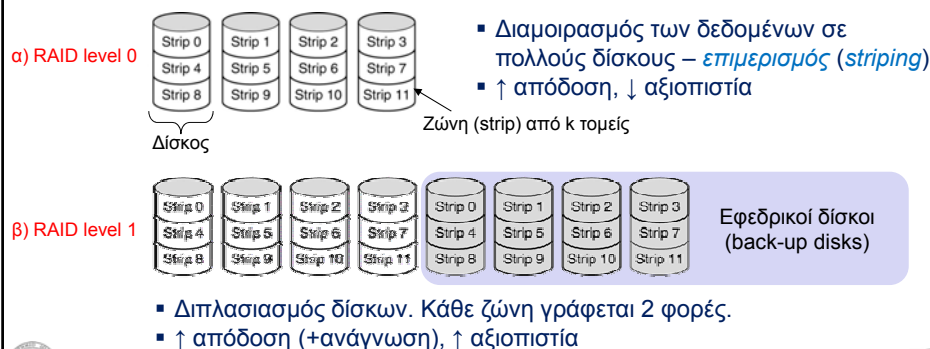
Μαγνητικοί Δίσκοι – Χαρακτηριστικά

- **Χρόνος αναζήτησης (seek time):** ο χρόνος τοποθέτησης της κεφαλής στην κατάλληλη ακτινική θέση (μερικά ms).
- **Λανθάνων χρόνος περιστροφής (rotational latency):** ο χρόνος τοποθέτησης του τομέα κάτω από την κεφαλή (μερικά ms).
 - Τυπικοί χρόνοι περιστροφής 3600, 5400, 7200, 10000, 15000 στροφές ανά λεπτό (rotations per minute – RPM).
- **Χρόνος μεταφοράς δεδομένων:** με ταχύτητες μεταφοράς της τάξης των εκατοντάδων MB/sec ο χρόνος μεταφοράς είναι της τάξης των εκατοστών του msec.
- Ο χρόνος αναζήτησης και ο λανθάνων χρόνος περιστροφής καθορίζουν την ταχύτητα απόκρισης του δίσκου.



Συστοιχίες RAID 1/3

- **Redundant Array of Inexpensive Disks – RAID (συστοιχία πλεοναζόντων φθηνών δίσκων):** οργάνωση των δίσκων για να βελτιωθεί η απόδοση ή/και η αξιοπιστία (Patterson 1988), σε αντιδιαστολή με τη χρήση ενός απλού μεγάλου ακριβού δίσκου (single large expensive disk – SLED).
- Έξι βασικοί τρόποι οργάνωσης (“επίπεδα” – levels): RAID 0-5.



Συστοιχίες RAID 2/3

γ) RAID level 2



- Οργάνωση πληροφορίας σε επίπεδο λέξης ή byte. Τα bits των λέξεων διαμοιράζονται σε πολλούς δίσκους.
- Bits ισοτιμίας – Κώδικας Hamming
- Απαιτεί πολλούς και συγχρονισμένους δίσκους.
- ↑ απόδοση, ↑ αξιοπιστία

δ) RAID level 3



- Οργάνωση πληροφορίας σε επίπεδο λέξης ή byte. Τα bits των λέξεων διαμοιράζονται σε πολλούς δίσκους.
- Ένα μόνο bit ισοτιμίας.
- Απαιτεί συγχρονισμένους δίσκους.
- ↑ απόδοση, ↑ αξιοπιστία



Οργάνωση Υπολογιστικών Συστημάτων – II

23

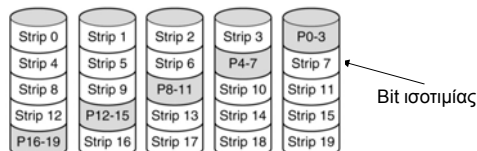
Συστοιχίες RAID 3/3

ε) RAID level 4



- Οργάνωση παρόμοια με το επίπεδο 0.
- Ένας δίσκος αποτελείται αποκλειστικά από ζώνες με bit ισοτιμίας.
- Ο υπολογισμός και η αποθήκευση των bit ισοτιμίας απαιτεί πολλαπλές αναγνώσεις/εγγραφές.
- ↓ απόδοση, ↑ αξιοπιστία

στ) RAID level 5



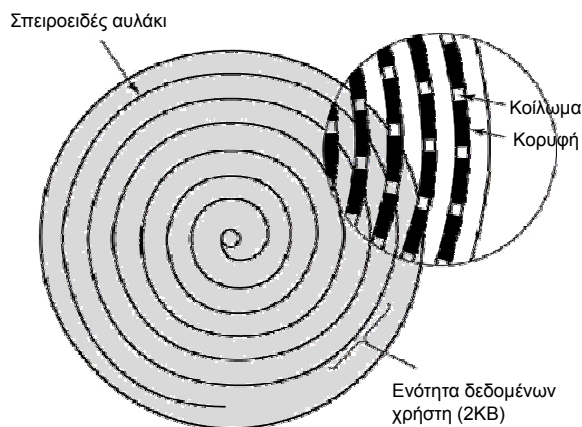
- Ομοιόμορφος διαμοιρασμός των bit ισοτιμίας σε όλους τους δίσκους με κυκλική σειρά. Μείωση φόρτου του δίσκου ισοτιμίας σε σχέση με το επίπεδο RAID-4.
- ↑ απόδοση, ↑ αξιοπιστία



Οργάνωση Υπολογιστικών Συστημάτων – II

24

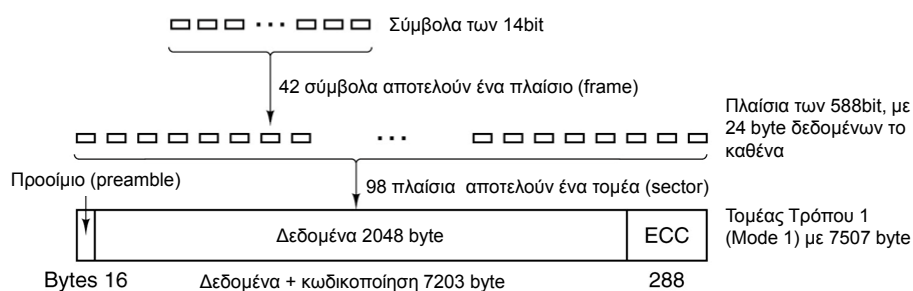
CD-ROM



- Ακτίνα 32mm.
- 22.188 περιστροφές → μήκος 5,6Km.
- Ανάγνωση: διοδικό laser χαμηλής ισχύος και με $\lambda=0.78\mu\text{m}$.
- Ταχύτητα περιστροφής 200 – 530 στροφές/min



CD-ROM – Οργάνωση



- Μηχανισμοί διόρθωσης σφαλμάτων σε επίπεδο συμβόλου, πλαισίου και τομέα.
- 98 πλαίσια των 588bit, δηλ. 7203 byte, για την αποθήκευση 2048 byte πληροφορίας. Συνεπώς η αποδοτικότητα είναι μόλις 28%.
- Δύο Τρόποι λειτουργίας.
- Χωρητικότητα 650MB (Τρόπος 1).
- Με ταχύτητα ανάγνωσης x32, ο ρυθμός μεταφοράς δεδομένων είναι 4MB/sec.



CD Εγγραφής και Επανεγγραφής

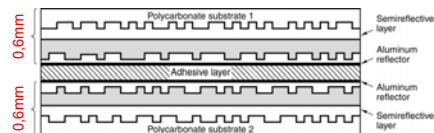
- **CD Εγγραφής (Recordable – CD-R):** Χρήση χρωστικής ύλης της οποίας η ανακλαστικότητα μεταβάλλεται με τη χρήση laser υψηλής ισχύος (8-16mW). Ανάγνωση με laser χαμηλής ισχύος (0,5mW).
- **CD Επανεγγραφής (Rewritable – CD-RW):** Χρήση κράματος μετάλλων που μπορεί να βρεθεί σε δύο καταστάσεις α) κρυσταλλική με υψηλή ανακλαστικότητα και β) άμορφη με χαμηλή ανακλαστικότητα.
 - Αρχική κατάσταση σε κρυσταλλική μορφή.
 - Μετάβαση σε άμορφη κατάσταση με χρήση laser υψηλής ισχύος (μέσω τήξης).
 - Επαναφορά στην κρυσταλλική κατάσταση με laser μέσης ισχύος.
 - Ανάγνωση με laser χαμηλής ισχύος.



DVD – Blu Ray

DVD

- Μικρότερα κοιλώματα από τα CD-ROM.
- Στενότερο σπειροειδές
- Κόκκινο laser ($\lambda=650\text{nm}$).
- Ταχύτητα: (x1) 1,4MB/sec, (x24) 33,24MB/sec.
- Πολλαπλών όψεων και στρώματων:
 - 1 όψη και 1 στρώμα → 4,7GB
 - 1 όψη και 2 στρώματα → 8,5GB
 - 2 όψεις και 1 στρώμα → 9,4GB
 - 2 όψεις και 2 στρώματα → 17GB



Blu Ray

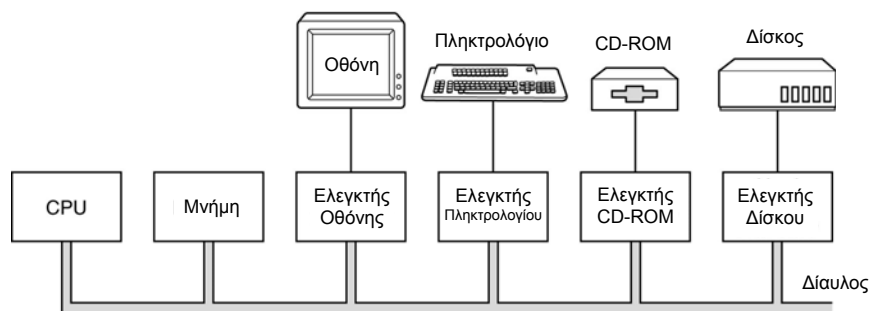
- Μπλε laser ($\lambda=405\text{nm}$).
- Ταχύτητα: (x1) 4,5MB/sec, (x12) 54MB/sec.
- Πολλαπλών όψεων και στρώματων:
 - 1 όψη και 1 στρώμα → 25GB
 - 1 όψη και 2 στρώματα → 50GB
 - 1 όψη και 3 στρώματα → 100GB
 - 1 όψη και 3 στρώματα → 128GB
 - 2 όψεις και 16 στρώματα → 1TB



ΕΙΣΟΔΟΣ / ΕΞΟΔΟΣ



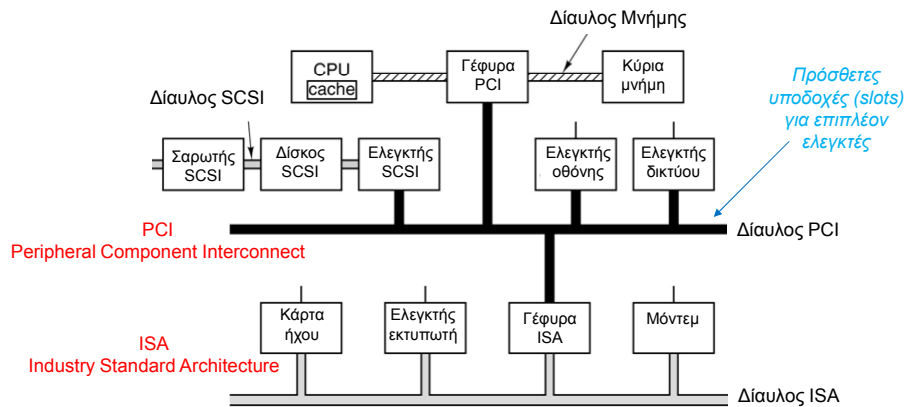
Τυπική Δομή Υπολογιστή – Δίαυλος



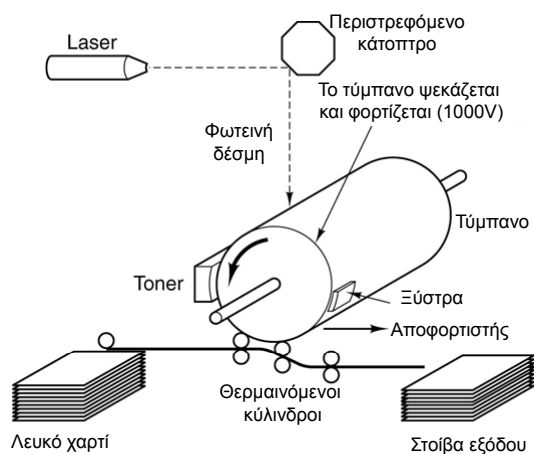
- Οι ελεγκτές καθορίζουν τη λειτουργία των αντίστοιχων συσκευών εισόδου/εξόδου και διαχειρίζονται την προσπέλαση του διαύλου.
- *Άμεση προσπέλαση μνήμης (direct memory access – DMA)*: ο ελεγκτής προσπελάνει τη μνήμη χωρίς ανάμιξη της CPU.
- Χρήση *διακοπών (interrupts)* για εξυπηρέτηση από CPU.
- Χρήση *διαιτητή διαύλου (bus arbiter)* για τον καθορισμό προτεραιοτήτων στο δίαυλο. Προτεραιότητα σε συσκευές I/O (*υπεξαίρεση κύκλων*).



Δομή Πολλαπλών Διαύλων



Εκτυπωτές Laser



Οθόνες LED

