

## Ενότητα 1<sup>η</sup>: Είσοδος / Έξοδος - Η απόδοση ενός συστήματος

**Σκοπός** Ο σκοπός της ενότητας αυτής είναι να παρουσιάσει τη μνήμη των υπολογιστών και τους τύπους της. Συγκεκριμένα θα επικεντρωθούμε στην κύρια μνήμη, την ιεραρχία, την οργάνωση και την απόδοση αυτής.

**Προσδοκώμενα Αποτελέσματα** Όταν θα έχετε μελετήσει την ενότητα, θα είστε σε θέση να:



περιγράφετε τη μνήμη των υπολογιστών και να αναφέρετε του τύπους της,



ιεραρχείτε τα είδη της μνήμης σε επίπεδα,



εξηγείτε τον τρόπο με τον οποίο οργανώνεται η κύρια μνήμη,



προσδιορίζετε τα κριτήρια για την απόδοση της κύριας μνήμης,



απαριθμείτε τις αρχιτεκτονικές για τη βελτίωση της απόδοσης της κύριας μνήμης,



προτείνετε τρόπους για την αποφυγή των συγκρούσεων των τραπεζών μνήμης.



μνήμη, μητρική πλακέτα, μπλοκ, κελί μνήμης, ιεραρχία μνήμης, κρυφή μνήμη, αρχή τοπικότητας, τοπικότητα χώρου, τοπικότητα χρόνου, ευρύτερη μνήμη, εναλλασσόμενη μνήμη, τράπεζα μνήμης, υπερ-τράπεζα μνήμης



## Η μνήμη υπολογιστών

Ένα από τα κύρια τμήματα κάθε υπολογιστή είναι η μνήμη. Με τον όρο μνήμη χαρακτηρίζουμε οποιαδήποτε μονάδα έχει τη δυνατότητα να αποθηκεύει δυαδικές πληροφορίες και να επιτρέπει την ασφαλή ανάκτησή τους. Στο εσωτερικό της κεντρικής μονάδας ενός υπολογιστή, μπορούμε να εντοπίσουμε μονάδες μνήμης, στον επεξεργαστή, στη μητρική πλακέτα ή σε κάρτες, καθώς και σε μαγνητικές ή οπτικές μονάδες (σκληροί δίσκοι κ.α.). Ορισμένες από τις μονάδες αυτές (κρυφή και κύρια μνήμη) είναι άμεσα προσπελάσιμες από τον επεξεργαστή (ΚΜΕ), ενώ ορισμένες άλλες (σκληρός δίσκος κ.ά.) είναι προσπελάσιμες μέσω μηχανισμών εισόδου-εξόδου. Επιπλέον μπορούν να χρησιμοποιηθούν μαγνητικές ή οπτικές μονάδες μνήμης, που βρίσκονται έξω από την κεντρική μονάδα του υπολογιστή, μέσω μηχανισμών εισόδου-εξόδου. Η μνήμη ενός υπολογιστή διακρίνεται συνήθως στην κύρια μνήμη, ή απλά μνήμη, και τη δευτερεύουσα ή βοηθητική μνήμη. Στην κύρια μνήμη περιλαμβάνεται η RAM (Random Access Memory) και η ROM (Read Only Memory) ενώ στις βοηθητικές μνήμες περιλαμβάνονται οι μαγνητικές και οι οπτικές μονάδες μνήμης.

### ■ Τυχαία και ακολουθιακή προσπέλαση

Οι μονάδες μνήμης μπορούν να διαχωριστούν σε μονάδες τυχαίας (ή άμεσης) προσπέλασης και μονάδες ακολουθιακής (ή σειριακής) προσπέλασης. Ο τρόπος προσπέλασης μιας μονάδας μνήμης καθορίζεται από τον τρόπο με τον οποίο μπορούν να αναζητηθούν και να ανακληθούν οι πληροφορίες που περιέχει η συγκεκριμένη μονάδα μνήμης. Ο χρόνος προσπέλασης (access time) μιας μνήμης είναι ο χρόνος που περνάει από τη στιγμή που επιλέγεται μια θέση μνήμης, μέχρι τη στιγμή που το περιεχόμενο της θέσης αυτής έχει διαβαστεί ή γραφτεί.

Μνήμες τυχαίας προσπέλασης: είναι κατασκευασμένες από ολοκληρωμένα κυκλώματα ημιαγωγών, δεν διαθέτουν κινούμενα μέρη και ο χρόνος προσπέλασης μιας θέσης μνήμης είναι πάντα ο ίδιος και είναι ανεξάρτητος από τη συγκεκριμένη θέση της μνήμης.

Μνήμες ακολουθιακής προσπέλασης: οι θέσεις μνήμης δεν είναι άμεσα διαθέσιμες και η πρόσβαση σε αυτές απαιτεί τη χρήση κινούμενων μερών (κεφαλές ανάγνωσης/εγγραφής). Για το λόγο αυτό, ο χρόνος προσπέλασης σε μια θέση μνήμης δεν είναι σταθερός, αλλά εξαρτάται από τη θέση της συγκεκριμένης θέσης μνήμης ως προς την κεφαλή.



**Η κύρια και η κρυφή μνήμη είναι τύπου τυχαίας προσπέλασης ενώ οι βοηθητικές μνήμες είναι τύπου ακολουθιακής προσπέλασης.**

## SRAM και DRAM

Τα ολοκληρωμένα κυκλώματα μονάδων τυχαίας προσπέλασης (RAM) διατίθενται σε δύο τύπους:

- Οι δυναμικές μνήμες **DRAM (Dynamic RAM)**
- Οι στατικές μνήμες **SRAM (Static RAM)**

➡ Μια στατική μνήμη περιλαμβάνει ουσιαστικά εσωτερικά flips-flops, που αποθηκεύουν δυαδικές πληροφορίες. Η αποθηκευμένη πληροφορία παραμένει έγκυρη (δεν αλλοιώνεται) για όσο παρέχεται ηλεκτρικό ρεύμα στη μονάδα μνήμης.

➡ Στη δυναμική μνήμη η δυαδική πληροφορία αποθηκεύεται με τη μορφή ηλεκτρικών φορτίων που εφαρμόζονται σε πυκνωτές. Το φορτίο που είναι αποθηκευμένο στους πυκνωτές έχει την τάση να ελαττώνεται με την πάροδο του χρόνου και για το λόγο αυτό θα πρέπει οι πυκνωτές περιοδικά να επαναφορτίζονται. Δηλαδή γίνεται ένα "φρεσκάρισμα" (refreshing) της δυναμικής μνήμης. Το φρεσκάρισμα πραγματοποιείται κάθε λίγα χιλιοστά του δευτερολέπτου με κυκλική σάρωση όλων των θέσεων μνήμης. Η δυναμική μνήμη σε σχέση με την στατική μνήμη χρησιμοποιεί λιγότερα ηλεκτρονικά στοιχεία και για το λόγο αυτό έχει μειωμένη κατανάλωση ισχύος, μεγαλύτερη χωρητικότητα αποθήκευσης ανά τσιπ μνήμης και μικρότερο κόστος. Αντίθετα η στατική μνήμη επειδή δεν χρειάζεται ανανέωση, είναι ταχύτερη από τη δυναμική.



Η κύρια μνήμη υλοποιείται με δυναμική μνήμη ενώ η κρυφή με στατική.



### Είδη DRAM στους Προσωπικούς Υπολογιστές

- Η **SDRAM (synchronous DRAM)** είναι ένα γενικό όνομα που δίνεται σε ορισμένες μνήμες τύπου DRAM οι οποίες μπορούν να συγχρονιστούν με την ταχύτητα του ρολογιού που δουλεύει ο επεξεργαστής. Η ταχύτητα μιας μνήμης SDRAM υπολογίζεται σε MHz και όχι σε nanosecond.
- Η **RDRAM (Rambus Dynamic Random Access Memory)** είναι ένα σχετικά νέα τεχνολογία μνήμης DRAM η οποία χρησιμοποιεί μια διαφορετική διάταξη των στοιχειωδών μονάδων αποθήκευσης που διαθέτουν τα ολοκληρωμένα κυκλώματα μνήμης. Η διάταξη αυτή με τη βοήθεια εκλεκτών και αρτηριών επιτρέπει στην μνήμη RDRAM να μεταφέρει μέχρι και 1.6 δισεκατομμύρια bytes το δευτερόλεπτο
- Η **DDR- SDRAM (Double Data Rate SDRAM)** είναι μια σύγχρονη δυναμική RAM η οποία θεωρητικά μπορεί να διπλασιάσει την ποσότητα των δεδομένων που μεταφέρονται σε μια δεδομένη συχνότητα λειτουργίας της. Αυτό το επιτυγχάνει εκμεταλλευόμενη και την άνοδο και την πτώση του σήματος κάθε κύκλου ρολογιού για τη μεταφορά δεδομένων.

## Πρόσκαιρη και μη πρόσκαιρη αποθήκευση

Οι μνήμες τα περιεχόμενα των οποίων χάνονται όταν πάψουμε να τις τροφοδοτούμε με ηλεκτρικό ρεύμα ονομάζονται "πρόσκαιρες" (*volatile*). Τα ολοκληρωμένα κυκλώματα των στατικών και δυναμικών RAM ανήκουν σε αυτή την κατηγορία αφού τα κύτταρα μνήμης τους χρειάζονται εξωτερική ηλεκτρική ισχύ για να διατηρήσουν

την αποθηκευμένη πληροφορία. Αντίθετα οι μη πρόσκαιρες μνήμες (nonvolatile) όπως είναι οι βοηθητικές μνήμες, διατηρούν τις πληροφορίες τους και χωρίς την ύπαρξη ηλεκτρικού ρεύματος. Αυτό συμβαίνει διότι στις μνήμες αυτές οι πληροφορίες αποθηκεύονται με μαγνητικό ή οπτικό και δεν επηρεάζονται από την διακοπή της ηλεκτρικής ισχύος. Άλλο παράδειγμα μη πρόσκαιρης μνήμης είναι η μνήμη μόνο-ανάγνωσης (ROM). Οι μη πρόσκαιρες μνήμες είναι απαραίτητες στους υπολογιστές για να μπορούμε να αποθηκεύουμε προγράμματα και δεδομένα που δεν χρησιμοποιούνται άμεσα αλλά θα τα χρειαστούμε στο μέλλον. Προγράμματα και δεδομένα που είναι απαραίτητα κατά τη διαδικασία εκκίνησης ενός υπολογιστή και δεν πρέπει να μεταβληθούν ποτέ αποθηκεύονται σε ROM. Ο υπολογιστής μπορεί να χρησιμοποιήσει τα περιεχόμενα της ROM όταν παρέχεται σε αυτόν ισχύ. Όταν ο υπολογιστής είναι σε λειτουργία και όταν χρειαστεί τα υπόλοιπα προγράμματα και δεδομένα που είναι αποθηκευμένα σε βοηθητικές μνήμες μπορούν να μεταφερθούν στην κύρια μνήμη (RAM) του υπολογιστή. Επίσης πριν την διακοπή της λειτουργίας του υπολογιστή τα περιεχόμενα της κύριας μνήμης, θα πρέπει να μεταφερθούν σε μια βοηθητική μνήμη εάν η πληροφορία αυτή πρέπει να φυλαχτεί.



### ΔΡΑΣΤΗΡΙΟΤΗΤΑ 1

- Πως διαχωρίζονται οι μονάδες μνήμης ανάλογα με τον τρόπο που γίνεται η προσπέλαση σε αυτές; Σε ποια κατηγορία ανήκουν η κρυφή και η κύρια μνήμη και σε ποια οι βοηθητικές μνήμες;
- Να εξηγήσετε τις είναι «ανανέωση» και γιατί είναι απαραίτητη στις μνήμες DRAM. Τι γνωρίζετε για τις μνήμες SRAM; Να συγκρίνετε τις απαντήσεις σας με την ενότητα «Η μνήμη υπολογιστών».
- Για περισσότερες λεπτομέρειες για τους τύπους μνήμης να ανατρέξετε στο βιβλίο του M. Morris Mano: «Ψηφιακή Σχεδίαση» και συγκεκριμένα στο 7<sup>ο</sup> κεφάλαιο: «Καταχωρητές, Μετρητές και Μονάδες Μνήμης».



## Η Ιεραρχία μνήμης



### ΔΡΑΣΤΗΡΙΟΤΗΤΑ 2

Η ταχύτητα με την οποία ο επεξεργαστής επεξεργάζεται τα δεδομένα είναι πολύ μεγαλύτερη από την ταχύτητα με την οποία η μνήμη μπορεί να τον τροφοδοτεί με δεδομένα. Τι συνέπεια έχει αυτό στην απόδοση του συστήματος; Να προτείνετε τρόπο με τον οποίο θα μπορούσε να αντιμετωπιστεί αυτό.



### ΑΠΑΝΤΗΣΗ ΔΡΑΣΤΗΡΙΟΤΗΤΑΣ 2

Η μικρή ταχύτητα της μνήμης σε σχέση με αυτή του επεξεργαστή έχει ως συνέπεια η μνήμη να αποτελεί έναν παράγοντα επιβράδυνσης της απόδοσης του επεξεργαστή και κατ' επέκταση ολόκληρου του συστήματος. Επομένως το ιδανικό θα ήταν να χρησιμοποιήσουμε μνήμη με μεγάλη χωρητικότητα αλλά και υψηλή ταχύτητα μεταφοράς δεδομένων. Μία τέτοια μνήμη όμως θα έχει πολύ υψηλό κόστος. Επομένως μία λύση θα ήταν να μπορούσαμε να χρησιμοποιούμε το είδος της μνήμης το οποίο είναι κατάλληλο σε κάθε περίπτωση, ανάλογα με τη χωρητικότητα και ταχύτητα που προσφέρει.



Για να αυξηθεί η συνολική απόδοση του υπολογιστικού συστήματος αναζητήθηκαν εναλλακτικοί τρόποι, έτσι ώστε η μνήμη να μην αποτελεί παράγοντα επιβράδυνσης του συστήματος. Ο κυριότερος από αυτούς τους τρόπους είναι η ιεράρχηση τμημάτων της μνήμης ανάλογα με τη ταχύτητα και τη χωρητικότητα που προσφέρουν. Στην κορυφή της ιεραρχίας έχουμε πολύ μικρές, γρήγορες αλλά ακριβές μνήμες, ενώ στη βάση έχουμε μνήμες με μεγάλο μέγεθος, αργές και φθηνές.

Σε μια τυπική ιεραρχία μνήμης (σχήμα 4.1.1), στο ανώτερο επίπεδο βρίσκονται οι καταχωρητές της ΚΜΕ, οι οποίοι χρησιμοποιούνται για την προσωρινή αποθήκευση εντολών και δεδομένων. Η ταχύτητα των καταχωρητών είναι πολύ υψηλή (ίση με τη ταχύτητα του επεξεργαστή), η χωρητικότητά τους είναι πολύ μικρή ενώ το κόστος κατασκευής των καταχωρητών είναι υψηλό.

Ακολουθεί η κρυφή μνήμη η οποία διαθέτει ταχύτητα μεγαλύτερη από εκείνη της κύριας μνήμης. Η κρυφή μνήμη αξιοποιεί την αρχή της τοπικότητας της αναφοράς και περιέχει αντίγραφα τμημάτων της κύριας μνήμης, με σκοπό τη γρήγορη τροφοδότηση του επεξεργαστή με τα δεδομένα που χρειάζεται. Συνήθως υλοποιείται με ψηφίδες στατικής μνήμης (SRAM).

Ακολουθεί η κύρια μνήμη, η οποία υλοποιείται με ψηφίδες δυναμικής μνήμης (DRAM), οι οποίες χαρακτηρίζονται από μεγαλύτερο χρόνο προσπέλασης και μικρότερο κόστος ανά μονάδα αποθηκευμένης πληροφορίας.

Οι μαγνητικοί και οι οπτικοί δίσκοι τοποθετούνται στα δύο επόμενα επίπεδα, ενώ στη βάση συναντούμε τις μονάδες μαγνητικών ταινιών, με κύρια χρήση τη μόνιμη και ασφαλή αποθήκευση μεγάλου όγκου δεδομένων, αλλά και τη διατήρηση αντιγράφων ασφαλείας. Λόγω των ηλεκτρομηχανικών διαδικασιών προσπέλασης στα δεδομένα, η ταχύτητα προσπέλασης είναι σχετικά χαμηλή. Οι μαγνητικοί δίσκοι προσφέρονται για την υλοποίηση του μηχανισμού της εικονικής μνήμης (στον οποίο θα αναφερθούμε στην 4η ενότητα του κεφαλαίου).



Σχήμα 4.1.1 - Η Ιεραρχία της μνήμης



Όσο κατεβαίνουμε τα επίπεδα της ιεραρχίας μνήμης αυξάνεται η χωρητικότητα της μνήμης και ο μέσος χρόνος προσπέλασης, ενώ μειώνεται το κόστος αποθήκευσης δεδομένων.



Ενώ τα χαρακτηριστικά της ταχύτητας και του κόστους των οπτικών δίσκων σε σχέση με τους μαγνητικούς δίσκους (σκληροί δίσκοι), είναι σύμφωνα με τη θέση τους στην ιεραρχία μνήμης, η χωρητικότητά τους έρχεται σε αντίθεση με αυτή.



Η ιεραρχία μνήμης είναι οργανωμένη έτσι ώστε να επιτρέπεται η επικοινωνία (αντιγραφή, μεταφορά δεδομένων) μόνον μεταξύ γειτονικών επιπέδων, δηλ. είτε με το προηγούμενο επίπεδο (αν έχει), είτε με το επόμενο του (αν έχει). Το ανώτερο επίπεδο είναι πιο κοντά στον επεξεργαστή και είναι μικρότερο και πιο γρήγορο από το κατώτερο. Η ελάχιστη ποσότητα δεδομένων η οποία μπορεί να μεταφερθεί από το ένα επίπεδο στο άλλο ονομάζεται μπλοκ. Στην απλούστερη περίπτωση ένα μπλοκ μπορεί να είναι μία λέξη.



Ιστορικά οι ΚΜΕ ήταν γρηγορότερες από τις μνήμες. Καθώς οι μνήμες βελτιώνονται, βελτιώνονται και οι ΚΜΕ, κι έτσι η ανισορροπία διατηρείται. Στην πράξη, όταν η ΚΜΕ δώσει μία αίτηση προς τη μνήμη, αναστέλλει τη λειτουργία της μέχρι η μνήμη να αποκριθεί. **Στους σύγχρονους υπολογιστές ανάμεσα στους καταχωρητές και την κύρια μνήμη τοποθετείται μια ενδιάμεση γρήγορη μνήμη που ονομάζεται κρυφή μνήμη. Η μνήμη αυτή βελτιώνει τη ταχύτητα επικοινωνίας του επεξεργαστή με την κύρια μνήμη.**



### **ΔΡΑΣΤΗΡΙΟΤΗΤΑ 3**

**Να κατατάξετε σε επίπεδα τα είδη της μνήμης, τοποθετώντας στο υψηλότερο επίπεδο το είδος με την μεγαλύτερη ταχύτητα. Να συγκρίνετε το σχήμα σας με το σχήμα 4.1.1.**



## Η μνήμη και η οργάνωσή της

Η μνήμη θεωρήθηκε σαν βασικό συστατικό της δομής ενός υπολογιστή, αρχικά από τον Turing και στη συνέχεια από τον Von Neumann, που έθεσε τις αρχές λειτουργίας των υπολογιστών. Η ύπαρξη της κύριας μνήμης σε έναν υπολογιστή είναι απαραίτητη γιατί ανεξάρτητα σε ποια βοηθητική μονάδα μνήμης βρίσκεται ένα πρόγραμμα (σκληρός δίσκος, δισκέτα, CD- ROM), για να εκτελεστεί αυτό το πρόγραμμα θα πρέπει να μεταφερθεί στην κύρια μνήμη και να παραμείνει εκεί για όλη τη διάρκεια εκτέλεσής του.

Η κύρια μνήμη ενός συστήματος υπολογιστή αποτελείται από ένα σύνολο κελιών ή κυψελίδων ή πιο απλά θέσεις μνήμης. Κάθε **κελί** ή θέση μπορεί να αποθηκεύσει ένα σταθερό αριθμό ψηφίων. Ο αριθμός αυτός συνήθως είναι 8 bits ή πολλαπλάσιό του και ονομάζεται **μήκος λέξης**. Κάθε κελί χαρακτηρίζεται από ένα μοναδικό αριθμό, ο οποίος προσδιορίζει τη διεύθυνσή του. Μέσω της διεύθυνσης αυτής ο επεξεργαστής και τα προγράμματα μπορούν να αναφέρονται στο περιεχόμενο του κελιού. Εάν μια μνήμη έχει  $m$  κελιά, τότε οι διευθύνσεις της είναι από το 0 μέχρι το  $m-1$  ή διαφορετικά, αν η διεύθυνση παριστάνεται με  $n$  δυαδικά ψηφία, οι διευθύνσεις της μνήμης κυμαίνονται από 0 έως  $2^n-1$ , ενώ ισχύει  $2^n = m$ . Οι διευθύνσεις των κελιών συνήθως εκφράζονται στο 16-δικό σύστημα. Το σχήμα 4.1.2 απεικονίζει την οργάνωση της μνήμης με τις λεπτομέρειες που προαναφέρθηκαν για τα κελιά, το μήκος τους και τις διευθύνσεις.



**Ο αριθμός που προσδιορίζει κάθε θέση μνήμης, η διεύθυνσή της, είναι μοναδικός και δεν αλλάζει, σ' αντίθεση με το περιεχόμενο των κελιών που είναι μεταβλητό.**

Το μέγεθος ενός κελιού μνήμης ονομάζεται **λέξη** και αποτελεί τη βασική μονάδα αναφοράς στην οργάνωση της μνήμης. Το μήκος της λέξης είναι συνήθως από 1 μέχρι και 8 bytes ή μεγαλύτερο, ανάλογα με τις δυνατότητες του επεξεργαστή. Η λέξη συνήθως καθορίζει και τη μικρότερη μονάδα μνήμης στην οποία μπορεί να έχει άμεση πρόσβαση ο επεξεργαστής.

## Οι λειτουργίες γραφής και ανάγνωσης

Οι δύο λειτουργίες που μπορεί να εκτελέσει μια μνήμη τυχαίας προσπέλασης είναι η γραφή και η ανάγνωση. Το σήμα γραφής κανονίζει τη λειτουργία μεταφοράς προς τη μνήμη και το σήμα ανάγνωσης τη λειτουργία μεταφοράς από τη μνήμη. Τα εσωτερικά κυκλώματα δέχονται ένα από αυτά τα σήματα ελέγχου και εκτελούν την επιθυμητή λειτουργία. Τα βήματα που ακολουθούνται για τη μεταφορά μιας λέξης στη μνήμη είναι τα ακόλουθα:

1. Μεταφορά της δυαδικής διεύθυνσης στις γραμμές διεύθυνσης.
2. Μεταφορά των bits δεδομένων, τα οποία θέλουμε να αποθηκευτούν στη μνήμη, στις γραμμές εισόδου δεδομένων.
3. Ενεργοποίηση της εισόδου ελέγχου γραφής.

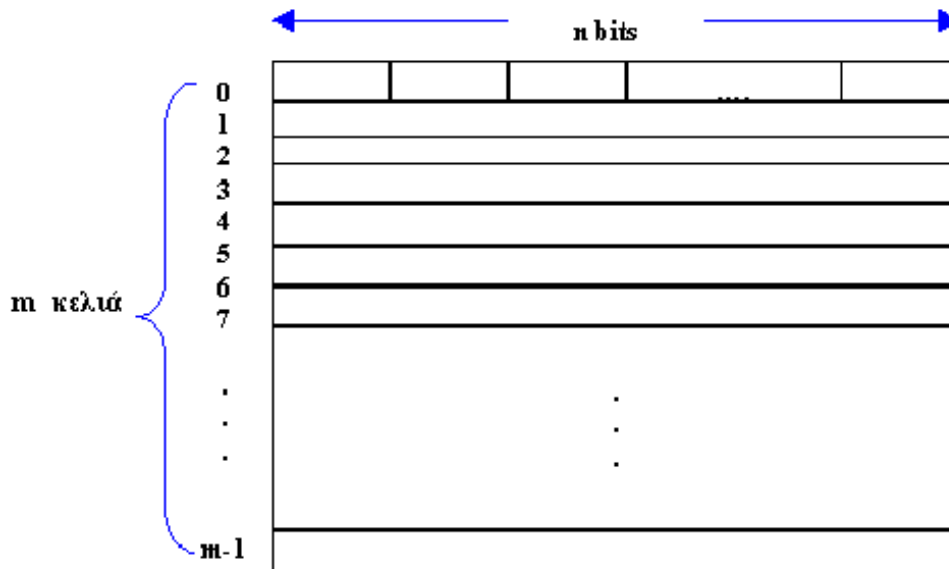
Η μονάδα μνήμης θα πάρει τότε τα bits από τις γραμμές δεδομένων και θα τις αποθηκεύσει στη λέξη που καθορίζουν οι γραμμές διεύθυνσης.



Τα βήματα που ακολουθούνται για τη μεταφορά μιας αποθηκευμένης λέξης έξω από τη μνήμη είναι τα ακόλουθα:

1. Μεταφορά της δυαδικής διεύθυνσης στις γραμμές διεύθυνσης.
2. Ενεργοποίηση της εισόδου ελέγχου ανάγνωσης.

Η μνήμη θα πάρει τότε τα bits από τη λέξη που έχει επιλεγεί από τη διεύθυνση και θα τα μεταφέρει στις γραμμές δεδομένων εξόδου. Το περιεχόμενο της επιλεγμένης λέξης δεν αλλάζει μετά την ανάγνωση.



4.1.2 – Η οργάνωση της κύριας μνήμης



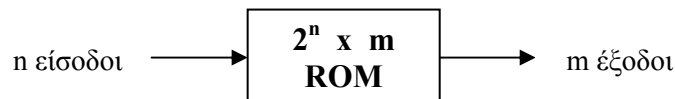
#### ΔΡΑΣΤΗΡΙΟΤΗΤΑ 4

Θυμάστε τον τρόπο με τον οποίο απεικονίζεται σχηματικά μια μνήμη ROM; Να σχεδιάσετε το συμβολικό διάγραμμα μιας μνήμης ROM. Για περισσότερες λεπτομέρειες για τη μνήμη ROM να ανατρέξετε στο βιβλίο του M. Morris Mano: «Ψηφιακή Σχεδίαση» και συγκεκριμένα στην ενότητα 5.7: «Μνήμη ανάγνωσης μόνο ROM».



#### ΑΠΑΝΤΗΣΗ ΔΡΑΣΤΗΡΙΟΤΗΤΑΣ 4

Η ROM είναι μία συσκευή μνήμης στην οποία αποθηκεύεται ένα σταθερό σύνολο δυαδικών πληροφοριών. Το σχήμα που ακολουθεί δείχνει το συμβολικό διάγραμμα μιας μνήμης ROM. Έχει  $n$  γραμμές εισόδου και  $m$  γραμμές εξόδου. Ο αριθμός των bits κάθε λέξης ισούται με τον αριθμό των γραμμών εξόδου, δηλαδή  $m$ . Ο αριθμός των δυνατών διαφορετικών διευθύνσεων για  $n$  μεταβλητές εισόδου είναι  $2^n$ . Μία ROM χαρακτηρίζεται από τον αριθμό των λέξεων  $2^n$  και τον αριθμό σε κάθε λέξη  $m$ .





## Η κύρια μνήμη και τα κριτήρια της απόδοσής της

Η κύρια μνήμη είναι το επόμενο επίπεδο στην ιεραρχία μνήμης μετά από την κρυφή μνήμη. Ικανοποιεί τις αιτήσεις της κρυφής μνήμης και αποτελεί το βασικό αποθηκευτικό χώρο για δεδομένα και εντολές, τα οποία είναι απαραίτητα για την εκτέλεση των προγραμμάτων, όταν ο υπολογιστής είναι σε λειτουργία. Η κύρια μνήμη επικοινωνεί με τις μονάδες Εισόδου/Εξόδου μέσω ενός κυκλώματος, μιας διεπαφής που εξασφαλίζει την ομαλή λειτουργία τους.



Στους πρώτους υπολογιστές η κύρια μνήμη είχε μικρή χωρητικότητα και πολύ υψηλό κόστος. Για να εκτελεστεί ένα μεγάλο πρόγραμμα έπρεπε να υποδιαιρεθεί από τον προγραμματιστή σε ανεξάρτητα τμήματα. Με την πάροδο του χρόνου το μέγεθος της κύριας μνήμης μεγάλωσε όπως μεγάλωσαν και τα προγράμματα. Η εμφάνιση των συστημάτων πολυπρογραμματισμού, επέβαλε στην κύρια μνήμη να μοιράζεται σε περισσότερα προγράμματα, γεγονός που δημιουργούσε νέα προβλήματα (όπως προβλήματα ασφάλειας ενός προγράμματος από τις επιδράσεις ενός άλλου). Η τεχνική της ιδεατής μνήμης και η ανταλλαγή εργασιών μεταξύ κύριας και δευτερεύουσας μνήμης έδωσαν μία λύση στα προβλήματα αυτά.

### Κριτήρια για την απόδοση της κύριας μνήμης

Σημαντικά κριτήρια για την απόδοση της κύριας μνήμης προκύπτουν από τον τρόπο οργάνωσης της κρυφής μνήμης και είναι:

- ο **χρόνος αναμονής**, ο οποίος είναι ο χρόνος που η λειτουργία της ΚΜΕ αναστέλλεται, ώστε να ολοκληρωθεί η λειτουργία της κρυφής μνήμης,
- το **εύρος ζώνης**, το οποίο είναι ο ρυθμός μεταφοράς δεδομένων από την κρυφή μνήμη προς τις μονάδες Εισόδου/Εξόδου.

Ο χρόνος αναμονής επηρεάζει την ποιότητα αποτυχίας της κρυφής μνήμης, ενώ το εύρος ζώνης είναι το πρωταρχικό μέλημα για τις μονάδες Εισόδου/Εξόδου γιατί επηρεάζει το χρόνο μεταφοράς των δεδομένων από και προς τις μονάδες αυτές.

Με την ευρεία χρησιμοποίηση των κρυφών μνημών δευτέρου επιπέδου και τα μεγάλα μεγέθη των μπλοκ τους, το εύρος ζώνης της κύριας μνήμης αυξάνεται σημαντικά.

Για τον χρόνο αναμονής χρησιμοποιούνται δύο μεγέθη:

- Ο **χρόνος πρόσβασης**, που είναι ο χρόνος μεταξύ μιας αίτησης ανάγνωσης και της στιγμής που φτάνει η επιθυμητή λέξη,
- Ο **χρόνος κύκλου**, που είναι ο ελάχιστος χρόνος μεταξύ αιτήσεων στη μνήμη.



Ο χρόνος κύκλου είναι πάντα μεγαλύτερος από τον χρόνο πρόσβασης γιατί οι γραμμές διευθύνσεων πρέπει να σταθεροποιηθούν προτού γίνει η επόμενη πρόσβαση.



## Αρχιτεκτονικές για τη βελτίωση της απόδοσης της κύριας μνήμης

Σημαντική παράμετρος για τις κρυφές μνήμες και κατ' επέκταση και την απόδοση της μνήμης, αποτελεί ο χαμηλός χρόνος αναμονής. Ωστόσο, σε γενικές γραμμές είναι ευκολότερο να βελτιώσεις το εύρος ζώνης της μνήμης με νέες αρχιτεκτονικές, παρά να μειώσεις την καθυστέρηση. Οι κρυφές μνήμες επωφελούνται από την αύξηση του εύρους ζώνης της κύριας μνήμης, αφού επιτρέπεται σε οποιουδήποτε μεγέθους μπλοκ να αυξήσει το μέγεθος του, χωρίς μεγάλη αύξηση στην ποινή αποτυχίας.

Έστω η περίπτωση μίας αποτυχίας, (όταν δεν βρεθεί το ζητούμενο μπλοκ στην κρυφή μνήμη). Ας δούμε πως υλοποιούνται αυτές οι αρχιτεκτονικές, υποθέτοντας ότι η απόδοση της κύριας μνήμης είναι:

- 4 κύκλοι για την αποστολή της διεύθυνσης
- 24 κύκλοι ο χρόνος πρόσβασης ανά λέξη
- 4 κύκλοι για την επιστροφή της λέξης

Διαθέτουμε μια κρυφή μνήμη με μπλοκ 4 λέξεων στην οποία θα έχουμε μια ποινή αποτυχίας:

$[4 \text{ λέξεις ανά μπλοκ} \times (4 \text{ κύκλοι για την αποστολή της διεύθυνσης} + \text{Χρόνος πρόσβασης ανά λέξη} + 4 \text{ κύκλοι για την επιστροφή της λέξης})] =$

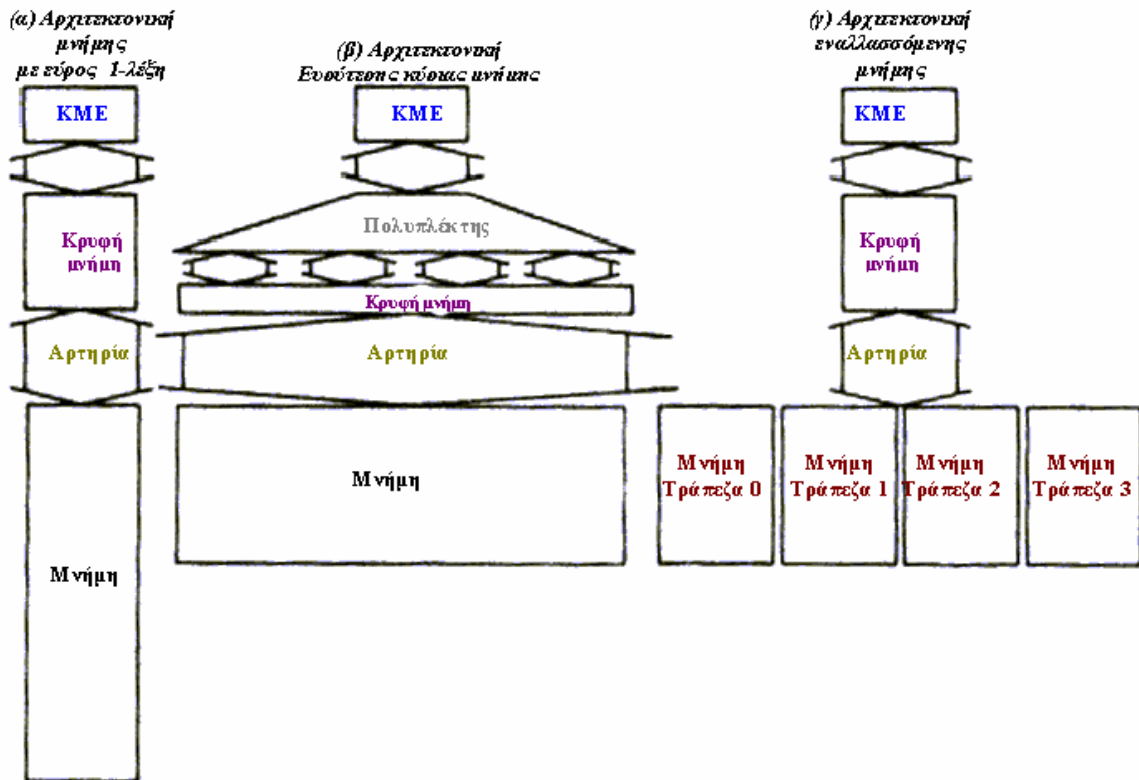
$4 \times (4 + 24 + 4) = 128 \text{ κύκλους, με ένα εύρος ζώνης } (16/128)=1/8 \text{ του byte ανά κύκλο ρολογιού.}$

Αξιοποιώντας τα δεδομένα του παραδείγματος, θα φανεί η βελτίωση της απόδοσης της κύριας μνήμης, με τη χρήση διαφόρων μεθόδων.

Οι τρεις σημαντικότερες αρχιτεκτονικές για τη βελτίωση της απόδοσης της Κύριας Μνήμης, είναι:

- **Η ευρύτερη κύρια μνήμη**
- **Η απλή εναλλασσόμενη μνήμη**
- **Η χρήση ανεξάρτητων τραπεζών μνήμης.**

Η απλούστερη προσέγγιση για την αύξηση του εύρους ζώνης είναι η ευρύτερη κύρια μνήμη. Στο σχήμα 4.1.3 παρουσιάζονται τρία παραδείγματα για το εύρος της αρτηρίας, μνήμης και εναλλασσόμενης μνήμης με απώτερο στόχο η μνήμη να αποκτήσει μεγαλύτερο εύρος ζώνης. Στο σχήμα (α) έχω την απλούστερη αρχιτεκτονική με εύρος 1 λέξη για την ΚΜΕ, την κρυφή μνήμη, την αρτηρία και την κύρια μνήμη. Στο σχήμα (β) παρουσιάζεται η αρχιτεκτονική μιας ευρύτερης κύριας μνήμης, ΚΜΕ, αρτηρίας και κρυφής μνήμης. Τέλος στο σχήμα (γ) παρουσιάζεται η αρχιτεκτονική μίας μικρού εύρους αρτηρίας, κρυφή μνήμη και μία εναλλασσόμενη κύρια μνήμη με 4 τράπεζες.



Σχήμα 4.1.3 - Παραδείγματα για εύρος αρτηρίας, μνήμης και εναλλασσόμενης μνήμης με απώτερο στόχο μνήμη με μεγαλύτερο εύρος ζώνης



## ΔΡΑΣΤΗΡΙΟΤΗΤΑ 5

Να περιγράψετε συνοπτικά τι είναι ο πολυπλέκτης και ποια η λειτουργία που επιτελεί. Να συγκρίνετε την απάντησή σας με την παράγραφο που ακολουθεί. Για περισσότερες λεπτομέρειες για τους πολυπλέκτες να ανατρέξετε στο βιβλίο του M. Morris Mano: «Ψηφιακή Σχεδίαση» και συγκεκριμένα στην ενότητα 5.6: «Πολυπλέκτες».



Ο πολυπλέκτης είναι γενικά μία συσκευή ή ένα κύκλωμα με δύο ή περισσότερες εισόδους και μία έξοδο, το οποίο υλοποιεί της λειτουργία της πολυπλεξίας. Η πολυπλεξία με τη σειρά της, αφορά:

- τη μεταφορά περισσότερων από δύο σημάτων σε μία γραμμή επικοινωνίας ή
- τη μεταφορά δεδομένων από χαμηλής ταχύτητας συσκευές αποθήκευσης, σε μία υψηλής ταχύτητας συσκευή, ή
- τέλος την επιλογή μιας από πολλές εισόδους δεδομένων και η μεταφορά της στη μοναδική έξοδο.



Ο υπολογιστής αποτελείται από πολλές μονάδες συνδεδεμένες μεταξύ τους με αρτηρίες δεδομένων, οι οποίες χρειάζονται συχνά σήματα ελέγχου για να ρυθμίζουν

τη δραστηριότητά τους. Τα σήματα ελέγχου δημιουργούνται από ένα τμήμα του υπολογιστή που ονομάζεται μονάδα ελέγχου. Η μονάδα ελέγχου είναι υπεύθυνη για τη μεταφορά, από τμήμα σε τμήμα, των σωστών δεδομένων την κατάλληλη στιγμή και για την επιβεβαίωση ότι λαμβάνουν χώρα οι κατάλληλες πράξεις πάνω στα δεδομένα. Σημαντικό τμήμα της μονάδας ελέγχου είναι το ρολόι (clock), το οποίο παράγει σήματα σε κανονικό ρυθμό. Αυτά τα σήματα χρησιμοποιούνται από τη μονάδα ελέγχου για να επιβεβαιωθεί ότι όλα τα τμήματα του υπολογιστή παραμένουν συγχρονισμένα μεταξύ τους και ότι στο κάθε τμήμα του υπολογιστή παρέχεται αρκετός χρόνος να ολοκληρώσει τη λειτουργία του.



## Η ευρύτερη κύρια μνήμη

Στα περισσότερα συστήματα υπολογιστών η κρυφή μνήμη είναι οργανωμένη σε μία πολυεπίπεδη διαβάθμιση, στην οποία συνήθως διακρίνουμε τα επίπεδα L1, L2. Οι κρυφές μνήμες πρώτου επιπέδου (L1) είναι οργανωμένες συνήθως με εύρος μίας λέξης γιατί οι περισσότερες προσβάσεις της ΚΜΕ είναι αυτού του εύρους. Τα συστήματα που δεν έχουν κρυφή μνήμη δευτέρου επιπέδου (L2) συνήθως σχεδιάζουν την κύρια μνήμη έτσι, ώστε αυτή να συμβαδίζει με το εύρος της κρυφής μνήμης. Διπλασιάζοντας ή τετραπλασιάζοντας το εύρος της κρυφής μνήμης και της κύριας μνήμης οδηγούμαστε στο διπλασιασμό ή τετραπλασιασμό του εύρους ζώνης της μνήμης. Έστω ότι η απόδοση της κύριας μνήμης είναι:

4 κύκλοι για την αποστολή της διεύθυνσης,

24 κύκλοι ο χρόνος πρόσβασης ανά λέξη και

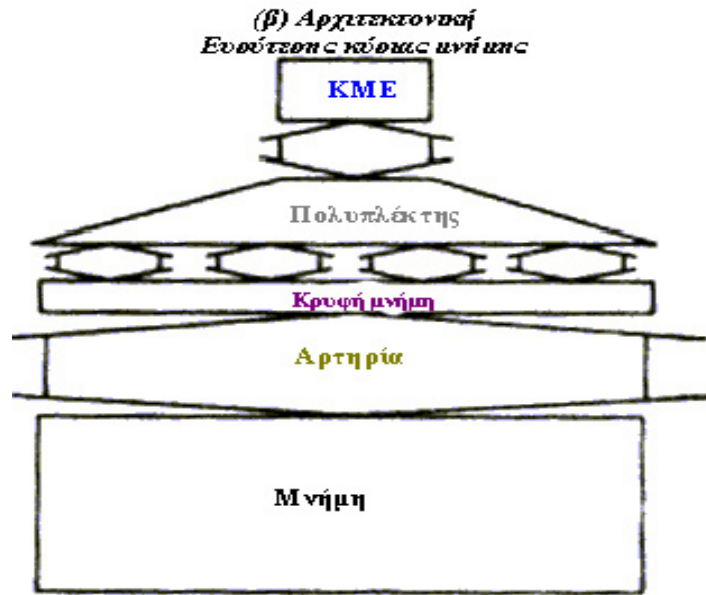
4 κύκλοι για την επιστροφή της λέξης.

Επιπλέον διαθέτουμε μια κρυφή μνήμη με μπλοκ 4 λέξεων. Με εύρος κύριας μνήμης 2 λέξεων η ποινή αποτυχίας στο παραπάνω παράδειγμα θα έπεφτε από  $4 \times 32 = 128$  σε  $2 \times 32 = 64$  κύκλους και το εύρος ζώνης γίνεται σε αυτή την περίπτωση  $(16/64=1/4)$  bytes ανά κύκλο ρολογιού. Με εύρος κύριας μνήμης 4 λέξεις η ποινή αποτυχίας γίνεται  $1 \times 32 = 32$  κύκλοι και το εύρος ζώνης είναι  $(16/32 = 1/2)$  bytes ανά κύκλο ρολογιού.

Τα συστήματα που δεν έχουν κρυφή μνήμη δευτέρου επιπέδου (L2) συνήθως σχεδιάζουν την κύρια μνήμη έτσι, ώστε αυτή να συμβαδίζει με το εύρος της κρυφής μνήμης.

Η αρτηρία μνήμης αποτελεί τον ενωτικό κρίκο μεταξύ της ΚΜΕ και της κύριας μνήμης στην αρχιτεκτονική της ευρύτερης κύριας μνήμης. Παρεμβάλλουμε το τμήμα από το σχήμα 4.1.3, που αναφέρεται στην ευρύτερη μνήμη (σχήμα 4.1.4).

Η ΚΜΕ θα συνεχίζει να προσπελαύνει την κρυφή μνήμη μία λέξη κάθε φορά και έτσι χρειάζεται ένας πολυπλέκτης μεταξύ κρυφής μνήμης και ΚΜΕ. Η τεχνική αυτή πρέπει να αποφεύγεται γιατί είναι χρονοβόρα. Οι κρυφές μνήμες δευτέρου επιπέδου (L2), βοηθούν αφού η πολυπλεξία μπορεί να γίνει μεταξύ πρώτου και δευτέρου επιπέδου και όχι μεταξύ της κρυφής μνήμης και της ΚΜΕ. Η ελάχιστη επέκταση που μπορούμε να κάνουμε είναι να διπλασιάσουμε ή να τετραπλασιάσουμε την κύρια μνήμη του ώστε να επωφεληθούμε αυτή την αύξηση του εύρους ζώνης.



Σχήμα 4.1.4 Αρχιτεκτονική ευρύτερης κύριας μνήμης

**!** Οι κύριες μνήμες με διόρθωση σφαλμάτων δυσκολεύονται στην αντιγραφή τμημάτων που περιέχουν προστατευμένα μπλοκ. Τα υπόλοιπα δεδομένα πρέπει να διαβαστούν έτσι ώστε να γίνει η ενημέρωση και η αποθήκευση παράλληλα με την εγγραφή των δεδομένων. Εάν η διόρθωση του σφάλματος γίνει σε όλο το εύρος της μνήμης, η ευρύτερη κύρια μνήμη θα αυξήσει τη συχνότητα ακολουθιών "ανάγνωση - τροποποίηση - εγγραφή", εξαιτίας των τμηματικών εγγραφών στα μπλοκ. Οι περισσότερες αρχιτεκτονικές ευρύτερης μνήμης διαθέτουν ξεχωριστή διόρθωση σφαλμάτων για κάθε 32 bits, όσο δηλαδή και το μέγεθος των εγγραφών.

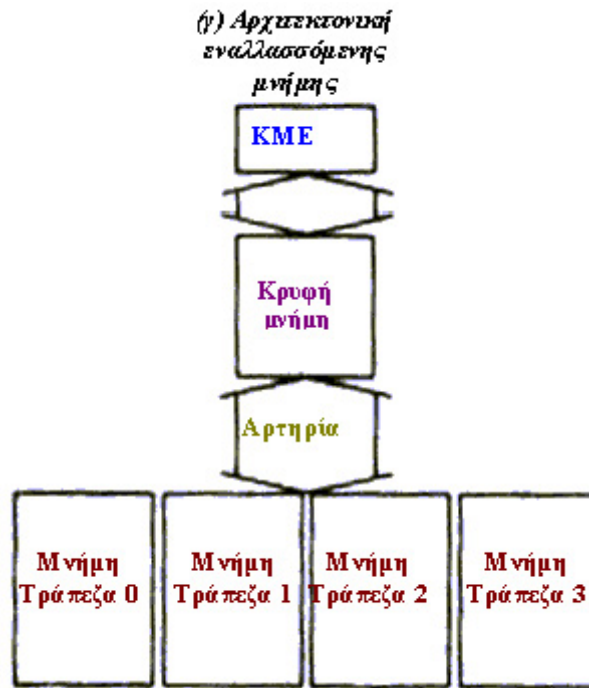


## Απλή εναλλασσόμενη μνήμη

Η αύξηση του εύρους ζώνης δεν αποτελεί τη μοναδική οδό για τη βελτίωση της απόδοσης της κύριας μνήμης. Εναλλακτικά χρησιμοποιούμε παράλληλα πολλές δυναμικές μνήμες (DRAMs) σε ένα σύστημα μνήμης.

Τα ολοκληρωμένα κυκλώματα της μνήμης μπορούν να οργανωθούν σε **τράπεζες** για να μπορούν να γράφουν ή να διαβάζουν περισσότερες από μία λέξεις τη φορά. Γενικά ο σκοπός της εναλλασσόμενης μνήμης είναι να προσπαθεί να εκμεταλλευτεί το εύρος ζώνης όλων των δυναμικών μνημών του συστήματος, σε αντίθεση με τα περισσότερα συστήματα μνήμης που ενεργοποιούν μόνο τις δυναμικές μνήμες που περιέχουν τις απαιτούμενες λέξεις.

Παραμβάλλουμε το τμήμα από το σχήμα 4.1.5, που αναφέρεται στην δομή της απλής εναλλασσόμενης μνήμης.



Σχήμα 4.1.5. Αρχιτεκτονική εναλλασσόμενης μνήμης

Οι τράπεζες έχουν συνήθως εύρος μίας λέξης, ώστε το μέγεθος της αρτηρίας και της κρυφής μνήμης να μην χρειάζεται αλλαγή, ενώ η αποστολή μιας διεύθυνσης σε έναν αριθμό από τράπεζες επιτρέπει σ' όλες να διαβάζουν ταυτόχρονα.



### Παράδειγμα

Αν υποθέσουμε ότι έχουμε μία κύρια μνήμη με απόδοση : 4 κύκλους για την αποστολή της διεύθυνσης, 24 κύκλους ο χρόνος πρόσβασης ανά λέξη και 4 κύκλους για την επιστροφή της λέξης. Επιπλέον διαθέτουμε μια κρυφή μνήμη με μπλοκ 4 λέξεων και μία ποινή αποτυχίας:  $4 \times (4 + 24 + 4) = 128$  κύκλους με ένα εύρος ζώνης  $(16/128) = 1/8 = 125$  του byte ανά κύκλο ρολογιού. Στέλνοντας μια διεύθυνση σε 4 τράπεζες έχουμε μια ποινή αποτυχίας  $4 + 24 + (4 \times 4) = 44$  κύκλους, δίνοντας ένα εύρος ζώνης  $16 / 44 = 0.4$  bytes ανά κύκλο ρολογιού. Αυτό συμβαίνει γιατί χρειάζονται 4 κύκλοι για την αποστολή της διεύθυνσης σε όλες τις τράπεζες (αφού αυτές μπορούν να την διαβάζουν ταυτόχρονα), 24 κύκλοι για την εύρεση της λέξης σε κάθε τράπεζα (αφού οι τράπεζες λειτουργούν παράλληλα) και κάθε μνήμη θέλει 4 κύκλους για να στείλει την λέξη πίσω. Επειδή όμως δεν μπορούν να στείλουν όλες τις λέξεις ταυτόχρονα χρειάζονται  $(4 \text{ κύκλοι}) \times (4 \text{ τράπεζες}) = 16$  κύκλους. 🏠



Η αντιστοίχιση των διευθύνσεων σε τράπεζες επηρεάζει τη συμπεριφορά του συστήματος μνήμης.



Υποθέτουμε ότι οι διευθύνσεις των τεσσάρων τραπεζών εναλλάσσονται σε επίπεδο λέξης. Δηλ. η τράπεζα 0 περιέχει όλες τις λέξεις των οποίων η (διεύθυνση modulo 4) είναι 0, η τράπεζα 1 τις λέξεις των οποίων η (διεύθυνση modulo 4) είναι 1 κ.ο.κ. Το σχήμα 4.1.6, δείχνει αυτή την εναλλαγή.

Διεύθυνση	Τράπεζα 0	Διεύθυνση	Τράπεζα 1
0		1	
4		5	
8		9	
12		13	

Διεύθυνση	Τράπεζα 2	Διεύθυνση	Τράπεζα 3
2		3	
6		7	
10		11	
14		15	

**Σχήμα 4.1.6** – Εναλλαγή διευθύνσεων των 4 τραπεζών σε επίπεδο λέξης. Εάν οι διευθύνσεις των τεσσάρων τραπεζών εναλλάσσονταν σε επίπεδο byte (δεδομένου ότι 4bytes=1 λέξη) κάθε μία από τις διευθύνσεις θα έπρεπε να πολλαπλασιαστεί επί 4.

Αυτή η αντιστοίχιση αναφέρεται ως *παράγοντας εναλλαγής*. Με τον όρο *εναλλασσόμενη μνήμη* συνήθως εννοούμε τράπεζες μνήμης όπου οι λέξεις εναλλάσσονται. Αυτή η εναλλαγή βελτιώνει την σειριακή (διαδοχική) προσπέλαση στην μνήμη. Μία αποτυχία ανάγνωσης της κρυφής μνήμης είναι ιδανική για μία μνήμη εναλλαγής λέξεων, καθώς οι λέξεις σε ένα μπλοκ διαβάζονται σειριακά.

Όσο αυξάνει η χωρητικότητα ανά Ο.Κ. (chip), υπάρχουν λιγότερα Ο.Κ. στο ίδιο μέγεθος σύστημα μνήμης, με αποτέλεσμα οι πολλαπλές τράπεζες να είναι πιο ακριβές. Ενώ σύμφωνα με τον κανόνα του Amdahl συνιστάται αύξηση της χωρητικότητας της μνήμης με την αύξηση της απόδοσης της ΚΜΕ, πολλοί κατασκευαστές προτιμούν μικρή μνήμη στο βασικό τους μοντέλο. Αυτός ο μειωμένος αριθμός των δυναμικών μνημών (DRAMs) είναι το κύριο μειονέκτημα των εναλλασσόμενων τραπεζών μνήμης. Ένα δεύτερο μειονέκτημα είναι η δυσκολία στην διερεύνηση της κύριας μνήμης.



## ΔΡΑΣΤΗΡΙΟΤΗΤΑ 6

- Να περιγράψετε συνοπτικά για ποιο σκοπό χρησιμοποιείται η εναλλασσόμενη μνήμη ως μέθοδος για τη βελτίωση της απόδοσης της μνήμης.
- Να εξηγήσετε τον τρόπο με τον οποίο λειτουργεί η σχεδίαση με ανεξάρτητες τράπεζες.





## ΑΠΑΝΤΗΣΗ ΔΡΑΣΤΗΡΙΟΤΗΤΑΣ 6

- Η εναλλασσόμενη μνήμη χρησιμοποιείται προκειμένου να εκμεταλλευτούμε την παράλληλη λειτουργία πολλών DRAM σε ένα σύστημα μνήμης. Τα ολοκληρωμένα κυκλώματα της μνήμης οργανώνονται σε τράπεζες για να μπορούν να γράφουν ή να διαβάζουν περισσότερες από μία λέξεις τη φορά. Γενικά, ο σκοπός της εναλλασσόμενης μνήμης είναι να εκμεταλλευτεί το εύρος ζώνης όλων των DRAM του συστήματος.
- Για να επιτευχθεί μεγαλύτερο εύρος ζώνης χρησιμοποιούνται τράπεζες μνήμης με εναλλασσόμενες σειριακές προσβάσεις. Μια γενίκευση της μεθόδου εναλλαγής είναι να επιτρέπονται πολλαπλές ανεξάρτητες προσβάσεις, όπου πολλαπλοί ελεγκτές μνήμης επιτρέπουν στις τράπεζες να λειτουργούν ανεξάρτητα. Κάθε τράπεζα χρειάζεται ξεχωριστές γραμμές διεύθυνσης και πιθανόν άλλη αρτηρία δεδομένων. Τέτοια σχεδίαση έχει νόημα να γίνει μόνο με τράπεζες μνήμης αλλιώς οι πολλαπλές αναγνώσεις θα εξυπηρετούνται από μια μνήμη και θα έχουν μόνο μικρό όφελος από αυτή τη δυνατότητα.



## Παράδειγμα

Θεωρούμε ότι έχουμε την παρακάτω περιγραφή για έναν υπολογιστή και την απόδοση της κρυφής μνήμης:

Μέγεθος μπλοκ = 1 λέξη

Εύρος αρτηρίας μνήμης = 1 λέξη

Ποσοστό αποτυχίας = 3%

Πρόσβαση μνήμης ανά εντολή = 1.2

Ποινή αποτυχίας κρυφής μνήμης = 32 κύκλοι

Μέσος όρος κύκλων ανά εντολή (εκτός των αποτυχιών της κρυφής μνήμης) = 2  
θεωρούμε πως αν αλλάξουμε το μέγεθος του μπλοκ σε 2 λέξεις το ποσοστό αποτυχίας πέφτει στο 2% και με μπλοκ 4 λέξεων έχουμε ποσοστό αποτυχίας 1%.

Ποια είναι η βελτίωση στην απόδοση μιας εναλλασσόμενης μνήμης 2 δρόμων και 4 δρόμων από όταν διπλασιάσουμε το μέγεθος της μνήμης και της αρτηρίας, υποθέτοντας ότι οι χρόνοι πρόσβασης είναι οι ακόλουθοι: 4 κύκλους για την αποστολή της διεύθυνσης, 24 κύκλους ο χρόνος πρόσβασης ανά λέξη και 4 κύκλους για την επιστροφή της λέξης;

## Απάντηση:

Οι κύκλοι ανά εντολή (CPI) για το βασικό υπολογιστή που χρησιμοποιεί μπλοκ 1 λέξης είναι:

(μέσος όρος κύκλων)+[(πρόσβαση μνήμης) x (% αποτυχίας) x (μέγεθος μπλοκ) x (ποινή αποτυχίας)]=  $2 + (1.2 * 3\% * 1 * (4+24+4)) = 3.15$

Στο παράδειγμα που εξετάζουμε δεν αλλάζει ο χρόνος κύκλου μηχανής ούτε η μέτρηση εντολών και έτσι μπορούμε να υπολογίσουμε τη βελτίωση της απόδοσης με βάση το CPI. Αυξάνοντας το μπλοκ της κρυφής μνήμης σε 2 λέξεις έχουμε τις παρακάτω επιλογές:

(Απευθείας εκτέλεση)


- 32-bit αρτηρία και μνήμη, μη εναλλασσόμενη =  $2 + [1,2 \times \% \text{αποτυχίας}_{\text{νέο}} \times (2 \text{ λέξεις} \times \text{ποινή αποτυχίας})] = 2 + [1,2 \times 2\% \times (2 \times 32)] = 3,54$
- 32-bit αρτηρία και μνήμη, εναλλασσόμενη =  $2 + [1,2 \times \% \text{αποτυχίας}_{\text{νέο}} \times [4 + 24 + (4 \text{ τράπεζες} \times 2 \text{ λέξεις})]] = [1,2 \times 2\% \times [4+24+(4 \times 2)]] = 2,86$
- 64-bit αρτηρία και μνήμη, μη εναλλασσόμενη =  $2 + (1,2 \times 2\% \times (1 \times 32)) = 2,77$

Διαπιστώνουμε ότι όταν διπλασιάσουμε το μέγεθος του μπλοκ επιβραδύνουμε την απευθείας εκτέλεση (3,54 έναντι 3,15) ενώ η εναλλασσόμενη ή η ευρύτερη μνήμη είναι 1,10 και 1,14 φορές ταχύτερη αντίστοιχα. Αν αυξήσουμε το μέγεθος του μπλοκ σε 4 λέξεις έχουμε:

- 32-bit αρτηρία και μνήμη, μη εναλλασσόμενη =  $2 + (1,2 \times 1\% \times (4 \times 32)) = 3,54$
- 32-bit αρτηρία και μνήμη, εναλλασσόμενη =  $2 + [1,2 \times \% \text{αποτυχίας}_{\text{νέο}} \times [4 + 24 + (4 \text{ τράπεζες} \times 4 \text{ λέξεις})]] = 2 + (1,2 \times 1\% \times (4+24+(4 \times 4))) = 2,53$
- 64-bit αρτηρία και μνήμη, μη εναλλασσόμενη =  $2 + (1,2 \times 1\% \times (2 \times 32)) = 2,77$

Παρατηρούμε ότι το μέγεθος του μπλοκ επιβαρύνει την απόδοση για την απλή περίπτωση παρόλο που η εναλλασσόμενη μνήμη των 32-bit είναι τώρα γρηγορότερη.

**(1,25 φορές γρηγορότερη έναντι 1,14 της ευρύτερης μνήμης & αρτηρίας)**

Η υποδιαίρεση αυτή μας δείχνει ότι η εναλλασσόμενη μνήμη είναι στην ουσία μια ευρύτερη μνήμη, μόνο που οι προσβάσεις στις τράπεζες ανά περιόδους αναγκάζονται να μοιραστούν εσωτερικούς πόρους (σε αυτήν την περίπτωση την αρτηρία). 



Ένα ερώτημα το οποίο προκύπτει για τις εναλλασσόμενες μνήμες είναι πόσες τράπεζες πρέπει να υπάρχουν. Ένα μέτρο που χρησιμοποιείται σε διανυσματικούς υπολογιστές είναι:

Αριθμός τραπεζών μνήμης  $\geq$  Αριθμός κύκλων για προσπέλαση μιας λέξης της τράπεζας.



## ΔΡΑΣΤΗΡΙΟΤΗΤΑ 7

Για ποιο λόγο θεωρείται ότι πρέπει να ισχύει η παραπάνω σχέση. Να αιτιολογήσετε την απάντησή σας δίνοντας ένα παράδειγμα.



## ΑΠΑΝΤΗΣΗ ΔΡΑΣΤΗΡΙΟΤΗΤΑΣ 7

Ο στόχος ενός συστήματος μνήμης είναι η μεταφορά πληροφορίας από μια νέα τράπεζα σε κάθε κύκλο ρολογιού, για ακολουθιακές προσβάσεις. Ας υποθέσουμε ότι ο αριθμός των τραπεζών μνήμης είναι μικρότερος από τον αριθμό των κύκλων για την προσπέλαση μιας λέξης της τράπεζας. Έστω ότι έχουμε 8 τράπεζες με χρόνο προσπέλασης 10 κύκλους ρολογιού. Μετά από 10 κύκλους ρολογιού η ΚΜΕ θα μπορεί να πάρει μία λέξη από την τράπεζα 0, και ενώ η τράπεζα 0 θα αρχίζει να φέρνει την επιθυμητή λέξη, η ΚΜΕ θα δέχεται τις επόμενες 7 λέξεις από τις υπόλοιπες 7 τράπεζες. Στον κύκλο ρολογιού 18 η ΚΜΕ θα πρέπει να περιμένει την τράπεζα 0 για να πάρει την επόμενη λέξη, μέχρι τον κύκλο ρολογιού 20. Για να αποφύγουμε λοιπόν αυτήν την αναμονή χρειαζόμαστε περισσότερες τράπεζες μνήμης από τους κύκλους προσπέλασης.



## Ανεξάρτητες τράπεζες μνήμης & Αποφυγή συγκρούσεων

Για να επιτευχθεί μεγαλύτερο εύρος ζώνης χρησιμοποιούνται τράπεζες μνήμης με εναλλασσόμενες σειριακές προσβάσεις. Μία γενίκευση της μεθόδου εναλλαγής είναι να επιτρέπονται πολλαπλές ανεξάρτητες προσβάσεις όπου πολλαπλοί ελεγκτές μνήμης θα επιτρέπουν στις τράπεζες να λειτουργούν ανεξάρτητα. Με έναν ελεγκτή μνήμης οι τράπεζες μπορούν να μοιράζονται τις γραμμές διευθύνσεων και έτσι η καθεμιά μπορεί να χρησιμοποιεί την πληροφορία που βρίσκεται στην αρτηρία μνήμης. Κάθε τράπεζα χρειάζεται ξεχωριστές γραμμές διευθυνσιοδότησης και πιθανόν άλλη αρτηρία δεδομένων. Για παράδειγμα, μια συσκευή Εισόδου/Εξόδου μπορεί να χρησιμοποιήσει έναν ελεγκτή και μία τράπεζα, η ανάγνωση της κρυφής μνήμης να χρησιμοποιήσει έναν άλλον και η εγγραφή της έναν τρίτο ταυτόχρονα. Τέτοια σχεδίαση έχει νόημα να γίνει μόνο με τράπεζες μνήμης, διαφορετικά οι πολλαπλές αναγνώσεις θα εξυπηρετούνται από μια μνήμη και θα έχουν μόνο μικρό όφελος από αυτή τη δυνατότητα. Όπως βλέπουμε ο όρος **τράπεζα μνήμης** έχει δύο διαφορετικούς ορισμούς. Γι' αυτό χρησιμοποιούμε τον όρο **υπέρ – τράπεζα** εννοώντας όλη την ενεργή μνήμη κατά τη μεταφορά ενός block και τον όρο **τράπεζα**

για το μέρος της υπέρ – μνήμης που είναι εναλλασσόμενης λέξης. Το σχήμα 4.1.7, δείχνει την σχέση αυτή.

Αριθμός Υπερ-τράπεζας	Μετατόπιση Υπερ-τράπεζας	
	Αριθμός τράπεζας	Μετατόπιση τράπεζας

Σχήμα 4. 1.7 – Η σχέση των υπερ-τραπεζών και των τραπεζών

### ■ Αποφυγή των συγκρούσεων των τραπεζών μνήμης

Όταν το σύστημα μνήμης είναι σχεδιασμένο να υποστηρίζει πολλαπλές ανεξάρτητες αιτήσεις (όπως στην περίπτωση των συσκευών Εισόδου/Εξόδου που έχουν απευθείας πρόσβαση στη μνήμη και μπορούν να γράφουν και να διαβάζουν δεδομένα από μη συνεχόμενες διευθύνσεις) η αποδοτικότητα του συστήματος εξαρτάται από την συχνότητα που οι ανεξάρτητες αιτήσεις πηγαίνουν σε διαφορετικές τράπεζες.

Η σειριακή προσπέλαση δεδομένων, όπως επίσης και οι προσβάσεις που οι διευθύνσεις διαφέρουν κατά περιττό αριθμό δουλεύουν καλά με την κλασσική εναλλασσόμενη μνήμη. Το πρόβλημα υπάρχει όταν η διαφορά των διευθύνσεων είναι άρτιος αριθμός γιατί τότε τα δεδομένα είναι πολύ πιθανόν να βρίσκονται στην ίδια τράπεζα. Μία λύση που προτάθηκε ήταν να μειωθεί το ποσοστό αυτό με την αύξηση του αριθμού των τραπεζών. Το πρόβλημα με αυτή τη λύση είναι ότι οι αναφορές στα δεδομένα της μνήμης δεν είναι τυχαίες και μπορούν να πάνε στην ίδια τράπεζα ανεξάρτητα από τον αριθμό των τραπεζών που υπάρχουν.



### Παράδειγμα:


Για παράδειγμα υποθέτουμε ότι έχουμε 128 τράπεζες μνήμης και εκτελούμε τον παρακάτω κώδικα:

```
int x[256][512]
for (j = 0; j < 512; j = j + 1)
for (i = 0; i < 256; i = i + 1)
x[i][j] = 2 * x[i][j];
```

### Απάντηση:

Με τον κώδικα αυτό διαβάζουμε όλα τα στοιχεία του πίνακα κατά στήλες, δηλαδή κατά σειρά  $x[0,0]$  ,  $x[1,0]$  ,  $x[2,0]$  κ.ο.κ. Τα στοιχεία αυτά όμως επειδή το 512 που είναι ο αριθμός των στηλών είναι άρτιο πολλαπλάσιο του 128 (που είναι ο αριθμός των τραπεζών), τα στοιχεία κάθε στήλης θα είναι στην ίδια τράπεζα. Πιο ειδικά, τα στοιχεία της στήλης 0 θα βρίσκονται στην τράπεζα 0, της στήλης 1 στην τράπεζα 1 κ.ο.κ. Έτσι το σύστημα θα κολλάει σε αποτυχίες της κρυφής μνήμης άσχετα με το πόσο πολύπλοκη ή σύγχρονη είναι η ΚΜΕ ή το σύστημα μνήμης.

Για την επίλυση του προβλήματος υπάρχουν λύσεις με λογισμικό αλλά και λύσεις με υλικό. Ο μεταγλωττιστής θα μπορούσε να κάνει βελτιστοποίηση της εναλλαγής του

βρόχου για να αποφύγει το πρόβλημα ή πιο απλά να επεκτείνει το μέγεθος του πίνακα ώστε αυτό να μην είναι δύναμη του 2 έτσι ώστε οι παραπάνω διευθύνσεις να πάνε σε διαφορετικές τράπεζες. 

Η αντιστοίχιση των διευθύνσεων σε μία θέση της τράπεζας μνήμης γίνεται με τις εξής εξισώσεις των τύπων 4.1, 4.2:

$$\text{Αριθμός τράπεζας} = (\text{Διεύθυνση}) \text{ MOD } (\text{Συνολικός αριθμός τραπεζών}) \quad (4.1)$$

$$\text{Διεύθυνση μέσα στην τράπεζα} = \text{Διεύθυνση} / \text{Συνολικό αριθμό τραπεζών} \quad (4.2)$$

Η λύση μέσω υλικού που δόθηκε για την μείωση των συγκρούσεων των τραπεζών μνήμης είναι ο περιορισμός το πλήθος των τραπεζών να είναι πρώτος αριθμός. Τέτοιοι αριθμοί απαιτούν συνήθως πολύπλοκους υπολογισμούς για να βρεθεί το υπόλοιπο και το πηλίκο αλλά υπάρχουν αρκετά σχήματα για να γίνει αυτός ο υπολογισμός γρήγορα. Σ' αυτή την περίπτωση η διαίρεση αντικαθίσταται από τον εξής υπολογισμό:

**Διεύθυνση μέσα στην τράπεζα = (Διεύθυνση) MOD (Αριθμός λέξεων στην τράπεζα)**  
Εφόσον ο αριθμός των λέξεων στην τράπεζα μνήμης είναι συνήθως δύναμη του 2 έχουμε αντικαταστήσει τη διαίρεση με την επιλογή του bit.

### Τράπεζα μνήμης

	<u>Σειριακά Εναλλασσόμενα</u>			<u>Εναλλασσόμενα κατά MOD</u>		
<i>Διεύθυνση στην τράπεζα</i>	0	1	2	0	1	2
0	0	1	2	0	16	8
1	3	4	5	9	1	17
2	6	7	8	18	10	2
3	9	10	11	3	19	11
4	12	13	14	12	4	20
5	15	16	17	21	13	5
6	18	19	20	6	22	14
7	21	22	23	15	7	23

**Σχήμα 4.1.8** – Παρουσιάζονται 3 στοιχεία μνήμης (0, 1, 2) το καθένα με 8 λέξεις δείχνοντας την κλασσική σειριακή εναλλασσόμενη αντιστοίχιση στα αριστερά και τη νέα αντιστοίχιση κατά modulo στα δεξιά.



Τα 3 στοιχεία μνήμης στα αριστερά του σχήματος 4.1.8, απαιτούν μια διαίρεση για την διευθυνσιοδότηση της λέξης μέσα στο στοιχείο ενώ η νέα αντιστοίχιση απαιτεί μόνο το υπόλοιπο μιας δύναμης του 2.



### Παράδειγμα:

Για τη διεύθυνση 5 έχουμε:

- Για την κλασσική αντιστοίχιση: Αριθμός τράπεζας =  $5 \text{ MOD } 3 = 2$   
Διεύθυνση μέσα στην τράπεζα =  $5 / 3 = 1$   
Άρα η διεύθυνση 5 αντιστοιχίζεται στην διεύθυνση 1 της τράπεζας (στοιχείου) 2.
- Για την νέα αντιστοίχιση: Αριθμός τράπεζας =  $5 \text{ MOD } 3 = 2$   
Διεύθυνση μέσα στην τράπεζα =  $5 \text{ MOD } 8 = 5$   
Άρα η διεύθυνση 5 αντιστοιχεί στην διεύθυνση 5 της τράπεζας (στοιχείου) 2.



### ΔΡΑΣΤΗΡΙΟΤΗΤΑ 8

Με ποιον τρόπο αποφεύγονται οι συγκρούσεις των τραπεζών μνήμης; Να συγκρίνετε την απάντησή σας με την παράγραφο: «Αποφυγή των συγκρούσεων των τραπεζών μνήμης».



### ΔΡΑΣΤΗΡΙΟΤΗΤΑ 9

Να βρεθεί η διεύθυνση 7, σε ένα σύστημα μνήμης που χρησιμοποιεί 4 τράπεζες μνήμης (0,1,2,3), η κάθε μία με 8 λέξεις, με:

- Κλασσική αντιστοίχιση
- Νέα αντιστοίχιση



### ΑΠΑΝΤΗΣΗ ΔΡΑΣΤΗΡΙΟΤΗΤΑΣ 9

Για την κλασσική αντιστοίχιση: Αριθμός τράπεζας =  $7 \text{ MOD } 4 = 3$

Διεύθυνση μέσα στην τράπεζα =  $7 / 4 = 1$

Άρα η διεύθυνση 7 αντιστοιχεί στην διεύθυνση 1 της τράπεζας 3.

Για την νέα αντιστοίχιση: Αριθμός τράπεζας =  $7 \text{ MOD } 4 = 3$

Διεύθυνση μέσα στην τράπεζα =  $7 \text{ MOD } 8 = 7$

Άρα η διεύθυνση 7 αντιστοιχεί στην διεύθυνση 7 της τράπεζας 3.

## Ανακεφαλαιώνοντας λοιπόν...



Η μνήμη γενικά είναι κάθε λειτουργική μονάδα ή μέσο στο οποίο φυλάσσονται δεδομένα και προς εκτέλεση εντολές με δυνατότητα ανάκτησης. Στην κύρια μνήμη περιλαμβάνεται η RAM και η ROM, ενώ στις βοηθητικές μνήμες περιλαμβάνονται οι μαγνητικές και οι οπτικές μονάδες μνήμης.



Οι μονάδες μνήμης μπορούν να χωριστούν σε μονάδες τυχαίας προσπέλασης και μονάδες ακολουθιακής προσπέλασης. Τα ολοκληρωμένα κυκλώματα μονάδων τυχαίας προσπέλασης (RAM) διατίθενται σε δύο τύπους: τις Δυναμικές μνήμες - DRAM και Στατικές μνήμες - SRAM.



Για να αυξήσουμε την συνολική απόδοση του υπολογιστικού συστήματος χρησιμοποιείται μία πολυεπίπεδη ιεραρχία μνήμης ανάλογα με τις επιδόσεις που προσφέρουν.



Η κύρια μνήμη ενός συστήματος υπολογιστή, αποτελείται από ένα σύνολο κελιών ή θέσεων. Κάθε κελί αποτελείται από ένα σταθερό αριθμό ψηφίων (bits), που συνήθως είναι 8 bits ή πολλαπλάσιό του.



Ο χρόνος αναμονής και το εύρος ζώνης αποτελούν τα πιο σημαντικά κριτήρια για την απόδοση της μνήμης. Η ευρύτερη κύρια μνήμη, η απλή εναλλασσόμενη μνήμη και η χρήση ανεξάρτητων τραπεζών μνήμης, αποτελούν τις τρεις βασικές τεχνικές βελτίωσης της απόδοσης της κύριας μνήμης.



Ο σκοπός της εναλλασσόμενης μνήμης είναι να εκμεταλλευτεί το εύρος ζώνης όλων των δυναμικών μνημών του συστήματος.



Για να επιτευχθεί μεγαλύτερο εύρος ζώνης χρησιμοποιούνται τράπεζες μνήμης με εναλλασσόμενες σειριακές προσβάσεις. Για να αποτρέψουμε τη σύγκρουση των τραπεζών της μνήμης δίνουμε λύσεις με χρήση λογισμικού και υλικού, όπως η απαίτηση το πλήθος των τραπεζών να είναι πρώτος αριθμός.