

---

# Προηγμένα Θέματα Αρχιτεκτονικής Συστημάτων

## Παράλληλη Επεξεργασία (μέρος 2<sup>ο</sup>)

Διδάσκων: Δημόκας Νικόλαος

# Περιεχόμενα

---

- **Παράλληλη Επεξεργασία**
  - Οργάνωση Πολλαπλών Επεξεργαστών
  - Συμμετρικοί Πολυεπεξεργαστές
  - Συνοχή Κρυφής Μνήμης και το Πρωτόκολλο MESI
  - Πολυνημάτωση και Chip Πολυεπεξεργαστών
  - Συστάδες
  - Μη Ομοιόμορφη Προσπέλαση Μνήμης
  - Υπολογισμοί Διανυσμάτων

# Πολυνημάτωση και Chip Πολυεπεξεργαστών - Εισαγωγή

---

- Η απόδοση ενός επεξεργαστή είναι ο ρυθμός με τον οποίο εκτελεί τις εντολές

$$\text{Ρυθμός MIPS} = f \times IPC$$

- $f$  είναι η συχνότητα του ρολογιού
  - $IPC$  (*Instructions Per Cycle*) είναι το μέσο πλήθος των εντολών οι οποίοι εκτελούνται ανά κύκλο
- Η αύξηση της απόδοσης κινείται σε δύο άξονες:
    1. Αύξηση της συχνότητας του ρολογιού
    2. Αύξηση του πλήθους των εκτελούμενων εντολών σε έναν κύκλο ρολογιού

# Πολυνημάτωση και Chip Πολυεπεξεργαστών - Εισαγωγή

---

- Οι σχεδιαστές έχουν αυξήσει την τιμή IPC χρησιμοποιώντας τη μέθοδο της διασωλήνωσης και έπειτα, χρησιμοποιώντας παράλληλες διασωληνώσεις σε μία υπερβαθμωτή αρχιτεκτονική
- Για να βελτιωθεί η ρυθμαπόδοση, οι σχεδιαστές δημιούργησαν ακόμη πιο πολύπλοκους μηχανισμούς, όπως είναι η *εκτέλεση ορισμένων εντολών με διαφορετική σειρά από εκείνη με την οποία εμφανίζονται μέσα στη ροή του προγράμματος*
- ❖ *Η προσέγγιση έχει φτάσει στα όριά της λόγω της πολυπλοκότητας και κατανάλωσης ισχύος*

# Πολυνημάτωση και Chip

## Πολυεπεξεργαστών - Εισαγωγή

---

- Μία εναλλακτική προσέγγιση είναι Πολυνημάτωση
- **Πολυνημάτωση:** η *ροή των εντολών διαιρείται σε διάφορες μικρότερες ροές*, γνωστές ως νήματα, με τέτοιο τρόπο, ώστε αυτά τα νήματα να **εκτελούνται παράλληλα**
- Η έννοια του νήματος το οποίο χρησιμοποιείται σε επεξεργαστές πολυνημάτωσης πιθανόν να είναι ή να μην είναι η ίδια με την έννοια των νημάτων λογισμικού την οποία συναντούμε σε ένα λειτουργικό σύστημα πολυπρογραμματισμού

# Έμμεση και Άμεση Πολυνημάτωση

---

- **Διεργασία:** Ένα **στιγμιότυπο προγράμματος** το οποίο **εκτελείται** σε έναν **υπολογιστή**. Η διεργασία ενσωματώνει δύο βασικά χαρακτηριστικά:

1. **Ιδιοκτησία πόρων:** Μία διεργασία **συμπεριλαμβάνει** ένα **χώρο ιδεατών διευθύνσεων** για να διατηρεί την εικόνα της διεργασίας. Η εικόνα διεργασίας είναι η συλλογή που περιέχει:

- το πρόγραμμα,
- τα δεδομένα,
- τη στοίβα και
- τα χαρακτηριστικά τα οποία ορίζουν τη διεργασία

Από χρονική στιγμή σε χρονική στιγμή, κατανέμεται στη διεργασία ο **έλεγχος των πόρων** όπως είναι η **κύρια μνήμη**, τα **κανάλια E/E**, οι **μονάδες E/E** και τα **αρχεία**.

# Έμμεση και Άμεση Πολυνημάτωση

---

- **Διεργασία:** Ένα **στιγμιότυπο προγράμματος** το οποίο **εκτελείται** σε έναν **υπολογιστή**. Η διεργασία ενσωματώνει δύο βασικά χαρακτηριστικά:
  2. **Χρονοδρομολόγηση/εκτέλεση:** Η **εκτέλεση** μίας διεργασίας είναι πιθανόν **διαστρωματωμένη** με την **εκτέλεση άλλων διεργασιών**. Η διεργασία έχει:
    - a. μία **κατάσταση εκτέλεσης** (Εκτελούμενη, Έτοιμη. κ.ο.κ)
    - b. μία **αναφορά προτεραιότητας**, και
    - c. έχει **χρονοδρομολογηθεί** και έχει λάβει προτεραιότητα από το λειτουργικό σύστημα.

# Έμμεση και Άμεση Πολυνημάτωση

---

- **Μεταγωγή διεργασίας:** Μία λειτουργία η οποία εναλλάσσει τον επεξεργαστή μεταξύ διαφορετικών διεργασιών αποθηκεύοντας όλα τα δεδομένα ελέγχου διεργασίας, τους καταχωρητές, αλλά και τις υπόλοιπες πληροφορίες της πρώτης και αντικαθιστώντας τις με τις αντίστοιχες πληροφορίες της δεύτερης



# Έμμεση και Άμεση Πολυνημάτωση

---

- **Νήμα:** Μία εκτελούμενη μονάδα εργασίας μέσα σε μία διεργασία. Ένα νήμα συμπεριλαμβάνει:
  - μετρητή προγράμματος,
  - δείκτη στοίβας,
  - και τη δική του περιοχή δεδομένων για τη στοίβα
- Ένα νήμα εκτελείται ακολουθιακά και υπάρχει δυνατότητα διακοπής του, ώστε ο επεξεργαστής να στραφεί σε ένα άλλο νήμα.
- **Μεταγωγή νήματος:** Η μεταφορά του ελέγχου του επεξεργαστή από ένα νήμα σε ένα άλλο, μέσα στην ίδια διεργασία

# Έμμεση και Άμεση Πολυνημάτωση

---

- Τα πολλαπλά **νήματα** εντός μίας διεργασίας **μοιράζονται** τους **ίδιους πόρους** → η **μεταγωγή νήματος** είναι λιγότερο χρονοβόρα από τη **μεταγωγή διεργασίας**
- Τα συστήματα εκτελούν ταυτόχρονα εντολές από διαφορετικά άμεσα νήματα:
  - είτε μέσω **φύλλωσης εντολών διαφορετικών νημάτων σε κοινόχρηστες διασωληνώσεις**
  - είτε με **παράλληλη εκτέλεση σε παράλληλες διασωληνώσεις**
- ❖ *Ένας επεξεργαστής πολλαπλών νημάτων θα πρέπει να έχει ένα ξεχωριστό μετρητή προγράμματος για κάθε νήμα που πρόκειται να εκτελεστεί ταυτόχρονα*

# Έμμεση και Άμεση Πολυνημάτωση

---

- Η έμμεση πολυνημάτωση αναφέρεται στην ταυτόχρονη εκτέλεση πολλαπλών νημάτων τα οποία εξάγονται από ένα ακολουθιακό πρόγραμμα.
  - Τα έμμεσα νήματα ορίζονται:
    - είτε στατικά από το μεταγλωττιστή,
    - είτε δυναμικά από το υλικό.
- ❖ Θα επικεντρωθούμε στην άμεση πολυνημάτωση

# Προσεγγίσεις Άμεσης Πολυνημάτωσης

---

Τέσσερις βασικές προσεγγίσεις για την πολυνημάτωση

- **Φυλλωμένη πολυνημάτωση:** Ο επεξεργαστής διαχειρίζεται **δύο ή περισσότερα νήματα ταυτόχρονα μεταφέροντας τον έλεγχο από το ένα στο άλλο σε κάθε κύκλο του ρολογιού**
  - Αν ένα νήμα **παρεμποδιστεί λόγω εξαρτήσεων στα δεδομένα ή λόγω λανθάνοντος χρόνου** της μνήμης, τότε ο επεξεργαστής **υπερπηδά** αυτό **το νήμα** και προχωρά στην εκτέλεση ενός άλλου
- **Οι εντολές από διαφορετικά νήματα δεν εκτελούνται ταυτόχρονα**
  - Ο επεξεργαστής μεταβαίνει από ένα νήμα σε ένα άλλο, χρησιμοποιώντας ένα διαφορετικό σύνολο από καταχωρητές

# Προσεγγίσεις Άμεσης Πολυνημάτωσης

---

Τέσσερις βασικές προσεγγίσεις για την πολυνημάτωση

- **Παρεμποδισμένη πολυνημάτωση:** Οι εντολές ενός νήματος **εκτελούνται διαδοχικά έως** ότου λάβει χώρα ένα **γεγονός** το οποίο είναι δυνατόν να **προκαλέσει καθυστέρηση**
  - Π.χ. η αστοχία της κρυφής μνήμης.
- Η μέθοδος είναι αποτελεσματική σε έναν επεξεργαστή ο οποίος εκδίδει εντολές σε σειρά και θα διακόψει τη διασωλήνωση αν υπάρξει ένα γεγονός καθυστέρησης
- Οι εντολές από διαφορετικά νήματα δεν εκτελούνται ταυτόχρονα

# Προσεγγίσεις Άμεσης Πολυνημάτωσης

---

- **Ταυτόχρονη πολυνημάτωση:** Οι εντολές εκδίδονται **ταυτόχρονα** από **πολλαπλά νήματα** προς **τις μονάδες εκτέλεσης** ενός **υπερβαθμωτού επεξεργαστή**
- Η προσέγγιση αυτή συνδυάζει τη δυνατότητα ευρείας έκδοσης εντολών των υπερβαθμωτών επεξεργαστών με τη χρήση πολλαπλών νημάτων.
- **Ταυτόχρονες εκτελέσεις εντολών** οι οποίες εκδίδονται από **διαφορετικά νήματα**

# Προσεγγίσεις Άμεσης Πολυνημάτωσης

---

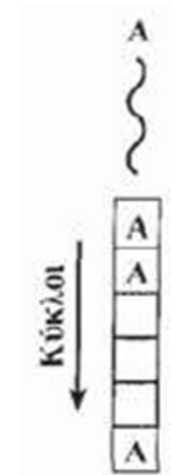
- **Πολυεπεξεργασία σε επίπεδο chip:** Σε αυτή την περίπτωση, ολόκληρος ο επεξεργαστής αντιγράφεται σε ένα chip και **κάθε επεξεργαστής διαχειρίζεται ξεχωριστά νήματα**
- Πλεονέκτημα αυτής της προσέγγισης είναι ότι δεν εξαρτάται από την αυξανόμενη πολυπλοκότητα σχεδίασης της διασωλήνωσης
- **Ταυτόχρονες εκτελέσεις εντολών** οι οποίες εκδίδονται από διαφορετικά νήματα

# Προσεγγίσεις Άμεσης Πολυνημάτωσης

Αρχιτεκτονικές διασωλήνωσης με πολυνημάτωση όταν χρησιμοποιείται ένας βαθμωτός επεξεργαστής:

- **Βαθμωτός απλής νημάτωσης:**

Πρόκειται για την **απλή διασωλήνωση** την οποία συναντάμε στις μηχανές CISC και RISC, **χωρίς πολυνημάτωση**



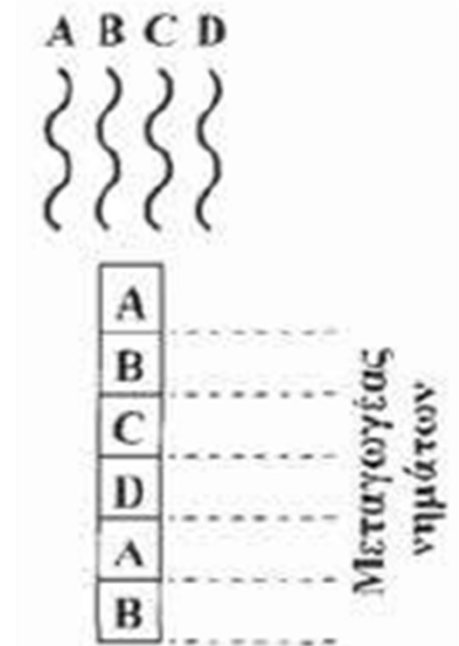
- Κατακόρυφα: Κύκλοι Ρολογιού
- Οριζόντια: Μέγιστο πλήθος εντολών που εκδίδονται εντός ενός κύκλου ρολογιού
- Ένα κενό κελί δείχνει μια αχρησιμοποίητη σχισμή διασωλήνωσης
- N: μία εντολή No-Op



# Προσεγγίσεις Άμεσης Πολυνημάτωσης

Αρχιτεκτονικές διασωλήνωσης με πολυνημάτωση όταν χρησιμοποιείται ένας βαθμωτός επεξεργαστής:

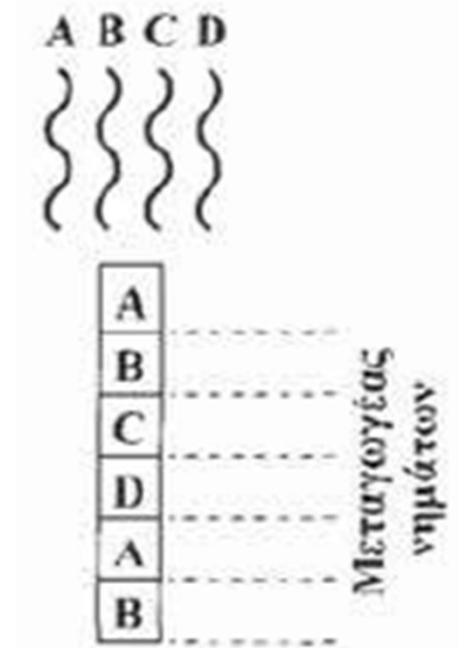
- **Βαθμωτός φυλωμένης πολυνημάτωσης:** Η απλούστερη στην υλοποίηση προσέγγιση πολυνημάτωση
- Μετάβαση από το ένα νήμα στο άλλο σε κάθε κύκλο του ρολογιού
- Τα στάδια της διασωλήνωσης διατηρούνται πλήρως απασχολημένα



# Προσεγγίσεις Άμεσης Πολυνημάτωσης

Αρχιτεκτονικές διασωλήνωσης με πολυνημάτωση όταν χρησιμοποιείται ένας βαθμωτός επεξεργαστής:

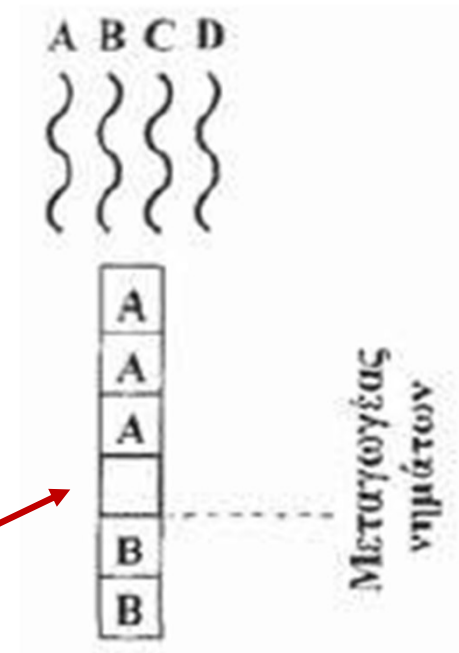
- Θεωρείται ότι δεν υπάρχουν εξαρτήσεις δεδομένων ή ελέγχου ανάμεσα στα νήματα
  - Για αυτό η μετάβαση γίνεται χωρίς καθυστέρηση (σε 0 κύκλους ρολογιού)



# Προσεγγίσεις Άμεσης Πολυνημάτωσης

Αρχιτεκτονικές διασωλήνωσης με πολυνημάτωση όταν χρησιμοποιείται ένας βαθμωτός επεξεργαστής:

- Βαθμωτός παρεμποδισμένης πολυνημάτωσης: Εκτελείται **ένα** απλό **νήμα** έως ότου **να εμφανιστεί** κάποιος **λανθάνων χρόνος**, ο οποίος θα σταματήσει τη διασωλήνωση.
  - Ο επεξεργαστής θα μεταβεί σε ένα άλλο νήμα.
- ❖ ο χρόνος που απαιτείται για μία μετάβαση από νήμα σε νήμα ισούται με ένα κύκλο.



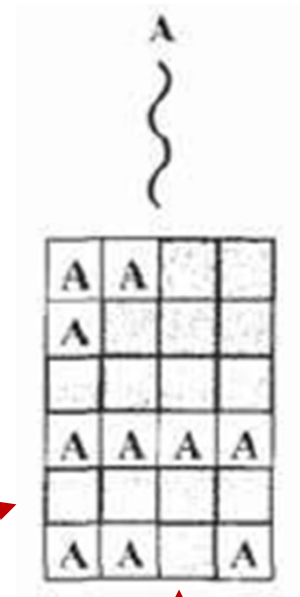
# Προσεγγίσεις Άμεσης Πολυνημάτωσης

---

- Η φυλλωμένη πολυνημάτωση φαίνεται να δίνει μεγαλύτερο βαθμό χρήσης του επεξεργαστή σε σύγκριση με την παρεμποδισμένη
- Αυτό επιτυγχάνεται θυσιάζοντας την απόδοση του απλού νήματος
  - Τα πολλαπλά νήματα ανταγωνίζονται για τους πόρους της κρυφής μνήμης
- Περισσότερες δυνατότητες για παράλληλη εκτέλεση διατίθενται αν ο επεξεργαστής έχει τη δυνατότητα να εκδίδει πολλαπλές εντολές ανά κύκλο

# Προσεγγίσεις Άμεσης Πολυνημάτωσης

- **Υπερβαθμωτή:** Πρόκειται για τη βασική υπερβαθμωτή προσέγγιση που **δεν χρησιμοποιεί πολυνημάτωση**
- Έως πρόσφατα, ήταν η πιο **ισχυρή προσέγγιση** όσον αφορά την **παροχή δυνατοτήτων παραλληλισμού** μέσα σε έναν επεξεργαστή.



Κατακόρυφη απώλεια

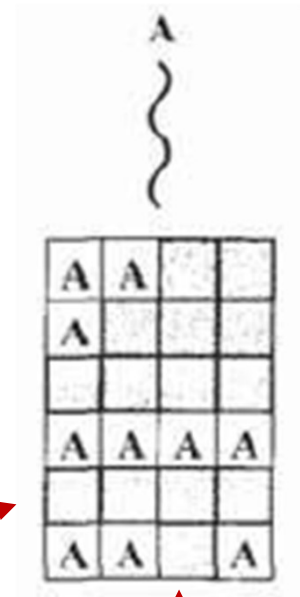
- **Απώλεια:** Όταν εκδίδονται λιγότερες εντολές από το μέγιστο πλήθος σε ένα κύκλο ρολογιού

Οριζόντια απώλεια

*Σε ένα κύκλο ρολογιού εκδίδονται εντολές ενός μόνο νήματος*

# Προσεγγίσεις Άμεσης Πολυνημάτωσης

- **Υπερβαθμωτή:** Πρόκειται για τη βασική υπερβαθμωτή προσέγγιση που **δεν χρησιμοποιεί πολυνημάτωση**
- Έως πρόσφατα, ήταν η πιο **ισχυρή προσέγγιση** όσον αφορά την **παροχή δυνατοτήτων παραλληλισμού** μέσα σε έναν επεξεργαστή.



Κατακόρυφη απώλεια

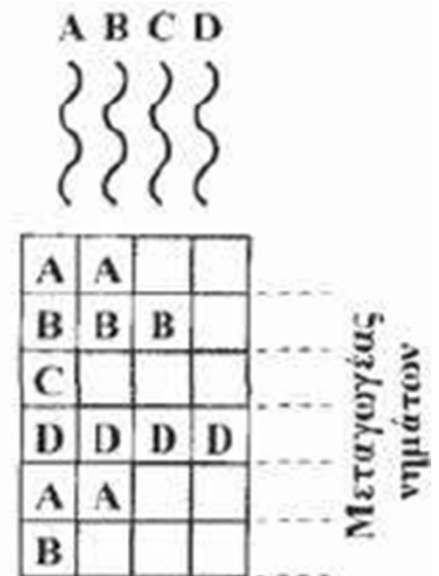
- **Απώλεια:** Όταν εκδίδονται λιγότερες εντολές από το μέγιστο πλήθος σε ένα κύκλο ρολογιού

Οριζόντια απώλεια

*Σε ένα κύκλο ρολογιού εκδίδονται εντολές ενός μόνο νήματος*

# Προσεγγίσεις Άμεσης Πολυνημάτωσης

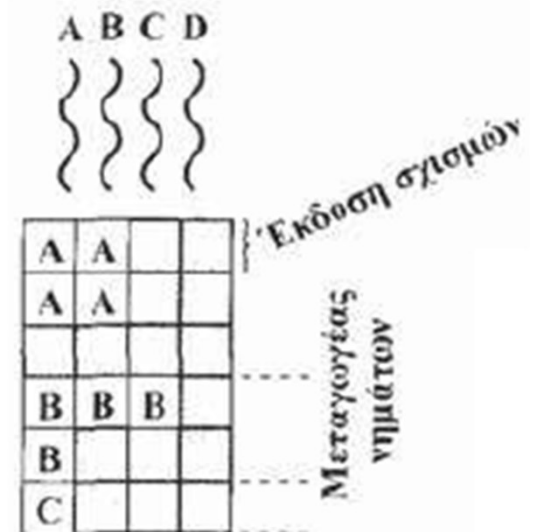
- Υπερβαθμωτή φυλλωμένη πολυνημάτωση: Σε ένα κύκλο, εκδίδονται όσο το δυνατόν περισσότερες εντολές από ένα απλό νήμα
- Έτσι εξαλείφονται οι πιθανές καθυστερήσεις που οφείλονται σε μεταβάσεις μεταξύ νημάτων
- Το πλήθος των εντολών οι οποίες εκδίδονται σε κάθε κύκλο, εξακολουθεί να είναι περιορισμένο λόγω των εξαρτήσεων οι οποίες υπάρχουν σε κάθε νήμα



*Σε ένα κύκλο ρολογιού εκδίδονται εντολές ενός μόνο νήματος*

# Προσεγγίσεις Άμεσης Πολυνημάτωσης

- Υπερβαθμωτή παρεμποδισμένη πολυνημάτωση: Οι εντολές ενός νήματος είναι πιθανό να εκδοθούν κατά τη διάρκεια οποιουδήποτε κύκλου και χρησιμοποιείται η παρεμποδισμένη πολυνημάτωση
- Παρεμποδισμένη: Εκτελούνται οι εντολές ενός νήματος μέχρι να εμφανιστεί κάποιος λανθάνων χρόνος



*Σε ένα κύκλο ρολογιού εκδίδονται εντολές ενός μόνο νήματος*



# Προσεγγίσεις Άμεσης Πολυνημάτωσης

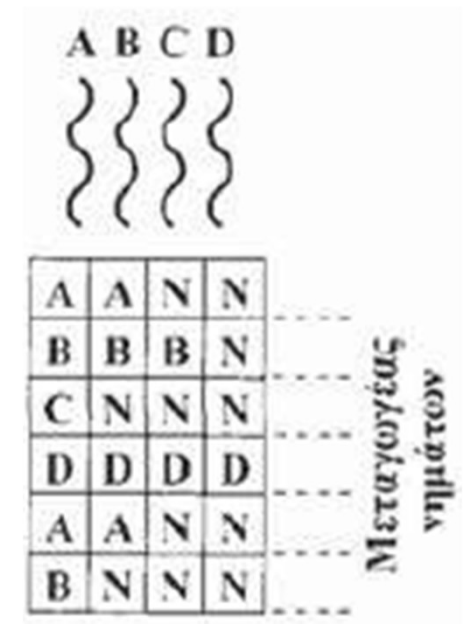
- **Λέξη εντολής πολύ μεγάλου μήκους (Very Long Instruction Word-VLIW):** Πολλαπλές εντολές τοποθετούνται σε μία λέξη. **Ο μεταγλωττιστής τοποθετεί στην ίδια λέξη τις λειτουργίες που μπορούν να εκτελεστούν παράλληλα**
- Αν δεν υπάρχουν εντολές για να εκτελεστούν παράλληλα τότε NO-OP (N) εντολή



*Σε ένα κύκλο ρολογιού εκδίδονται εντολές ενός μόνο νήματος*

# Προσεγγίσεις Άμεσης Πολυνημάτωσης

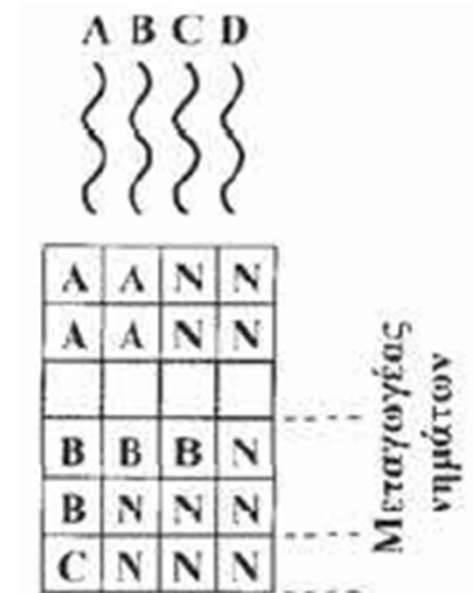
- **Φυλλωμένη πολυνημάτωση VLIW:** Αυτή η προσέγγιση θα πρέπει να παρέχει **παρόμοια αποδοτικότητα** με εκείνη που παρέχει η φυλλωμένη πολυνημάτωση σε μία υπερβαθμωτή αρχιτεκτονική



*Σε ένα κύκλο ρολογιού εκδίδονται εντολές ενός μόνο νήματος*

# Προσεγγίσεις Άμεσης Πολυνημάτωσης

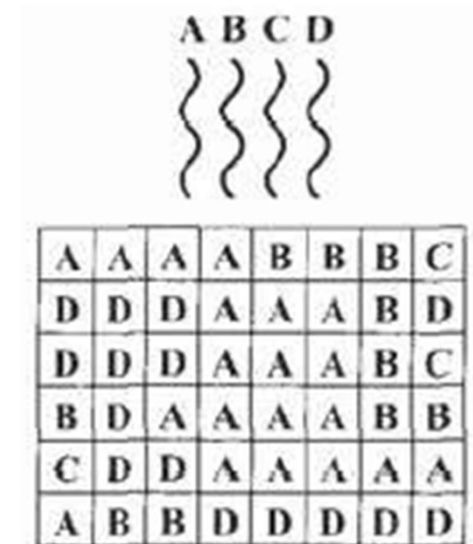
- **Παρεμποδισμένη πολυνημάτωση VLIW:**  
Αυτή η προσέγγιση θα πρέπει να παρέχει **παρόμοια αποδοτικότητα** με εκείνη που παρέχει η **παρεμποδισμένη πολυνημάτωση σε μία υπερβαθμωτή αρχιτεκτονική**



*Σε ένα κύκλο ρολογιού εκδίδονται εντολές ενός μόνο νήματος*

# Προσεγγίσεις Άμεσης Πολυνημάτωσης

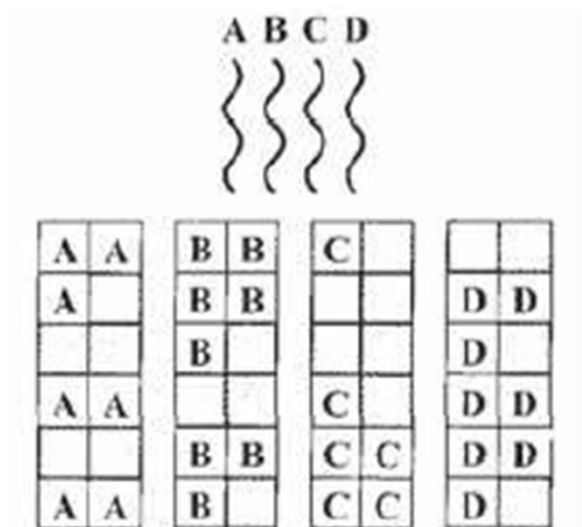
- **Ταυτόχρονη πολυνημάτωση:** Αν ένα νήμα έχει **μεγάλο βαθμό παραλληλισμού** σε επίπεδο εντολής **μπορεί να γεμίσει τις οριζόντιες σχισμές σε μερικούς κύκλους**
- Σε κάποιους κύκλους, μπορεί να εκδοθούν οι εντολές από δύο ή περισσότερα νήματα.



*Σε ένα κύκλο ρολογιού επιτρέπεται η ταυτόχρονη εκτέλεση πολλαπλών νημάτων*

# Προσεγγίσεις Άμεσης Πολυνημάτωσης

- **Chip Πολυεπεξεργαστών** (πολλαπλών πυρήνων): Σε κάθε επεξεργαστή εκχωρείται ένα νήμα από το οποίο έχει τη δυνατότητα έκδοσης έως και δύο εντολών ανά κύκλο



*Σε ένα κύκλο ρολογιού επιτρέπεται η ταυτόχρονη εκτέλεση πολλαπλών νημάτων*

# Προσεγγίσεις Άμεσης Πολυνημάτωσης

- Σύγκριση Ταυτόχρονης Πολυνημάτωσης με Chip Πολυεπεξεργαστών
- Ένα **Chip Πολυεπεξεργαστών**, που έχει την ίδια δυνατότητα έκδοσης εντολών με έναν επεξεργαστή Ταυτόχρονης Πολυνημάτωσης, *δεν μπορεί να επιτύχει τον ίδιο βαθμό παραλληλισμού σε επίπεδο εντολής* γιατί
  - ο πολυεπεξεργαστής δεν έχει τη δυνατότητα να επικαλύψει τον λανθάνοντα χρόνο εκδίδοντας εντολές από άλλα νήματα



A	A	A	A	B	B	B	C
D	D	D	A	A	A	B	D
D	D	D	A	A	A	B	C
B	D	A	A	A	A	B	B
C	D	D	A	A	A	A	A
A	B	B	D	D	D	D	D



A	A	B	B	C				
A		B	B				D	D
		B					D	
A	A			C			D	D
		B	B	C	C		D	D
A	A	B		C	C		D	

# Συστάδες

---

- Η **Συσταδοποίηση** αποτελεί μία **εναλλακτική λύση** σε σχέση με τη **συμμετρική πολυεπεξεργασία**, ως προσέγγιση η οποία **παρέχει υψηλή απόδοση και διαθεσιμότητα**
  - Χρησιμοποιείται σε Εξυπηρετητές (Servers)
- **Συστάδα**: Ένα σύνολο **διασυνδεδεμένων, αυτόνομων υπολογιστών**, οι οποίοι εργάζονται μαζί ως **ενοποιημένη υπολογιστική μηχανή**
  - Κάθε υπολογιστής μίας συστάδας, αναφέρεται ως κόμβος

# Συστάδες

---

## Πλεονεκτήματα της Συσταδοποίησης

- 1. Απόλυτη κλιμάκωση:** Μπορούν να δημιουργηθούν *μεγάλες συστάδες* οι οποίες *ξεπερνούν την ισχύ* ακόμη και *των μεγαλύτερων αυτόνομων μηχανών*.
  - Μία συστάδα μπορεί να περιέχει δεκάδες, εκατοντάδες, ακόμη και χιλιάδες μηχανές, κάθε μία από τις οποίες είναι ένας πολυεπεξεργαστής
- 2. Επαυξητική κλιμάκωση:** Μία συστάδα διαμορφώνεται με τέτοιο τρόπο ώστε να είναι *εφικτή η προσθήκη νέων συστημάτων* σε αυτή, με μικρές επαυξήσεις.



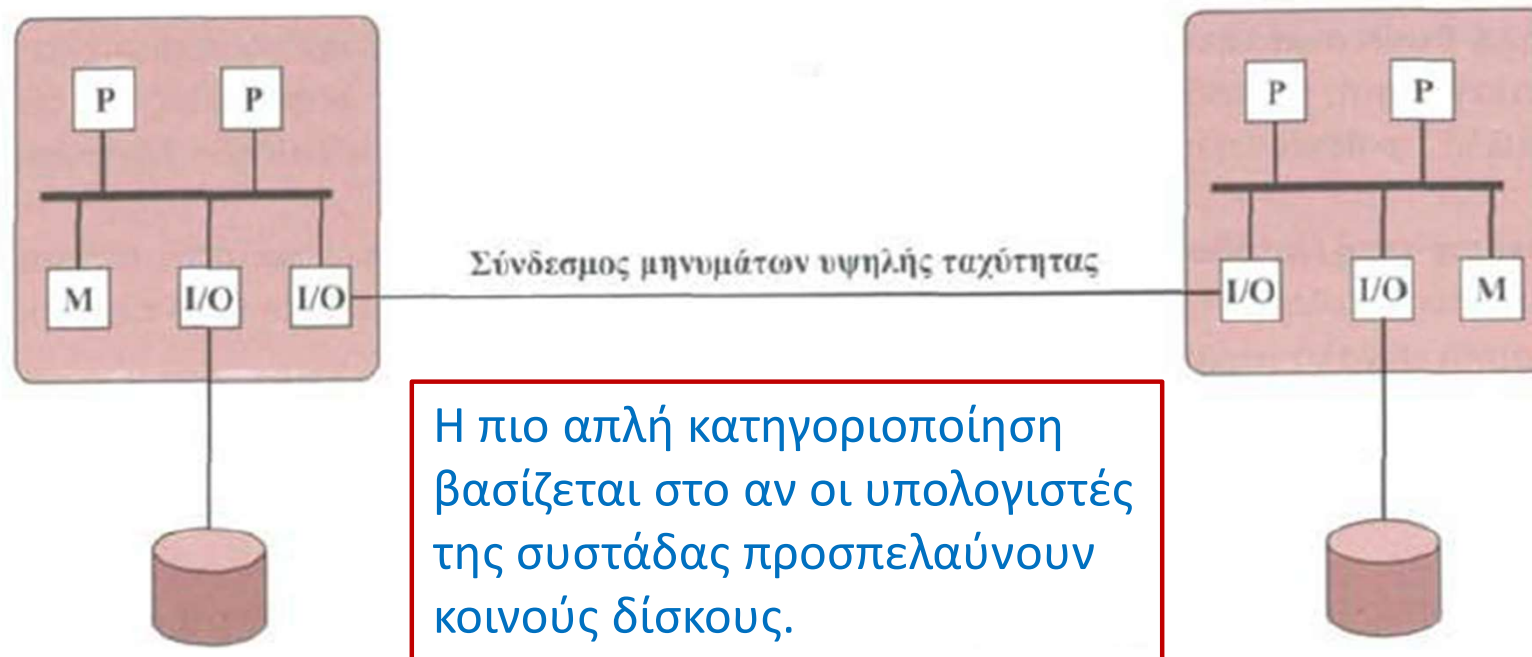
# Συστάδες

---

## *Πλεονεκτήματα της Συσταδοποίησης*

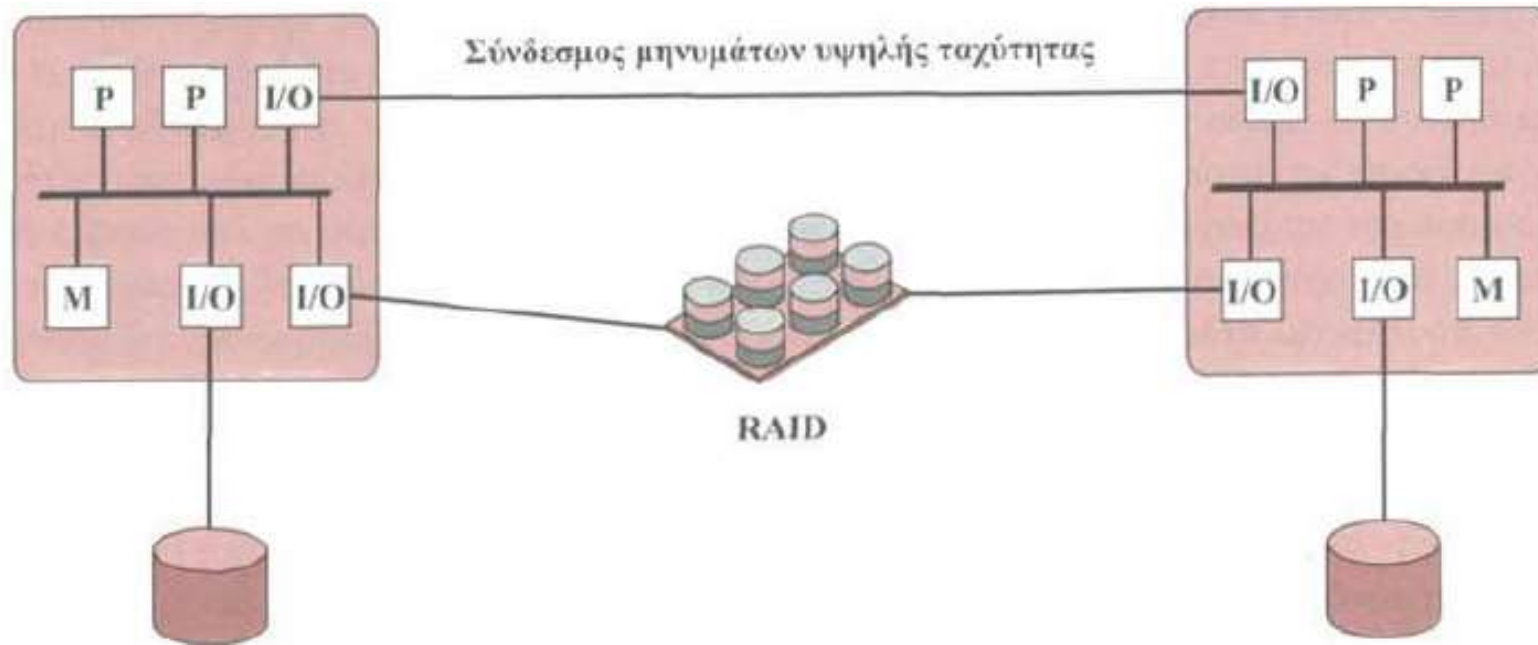
- 3. Υψηλή διαθεσιμότητα:** Λόγω του γεγονότος ότι κάθε κόμβος σε μία συστάδα αποτελεί αυτόνομο υπολογιστή, *η βλάβη ενός κόμβου δεν σημαίνει την απώλεια της λειτουργίας της συστάδας*
- 4. Ανώτερη τιμή/απόδοση:** Μπορούμε *να φτιάξουμε μία συστάδα* με ίση ή μεγαλύτερη υπολογιστική ισχύ σε σύγκριση με μία απλή μεγάλη μηχανή, χρησιμοποιώντας *απλούς υπολογιστές με πολύ χαμηλό κόστος*

# Συστάδες - Διατάξεις



- Μία συστάδα **χωρίς κοινόχρηστο δίσκο** αποτελούμενη από δύο κόμβους, όπου η μοναδική διασύνδεση είναι ένας σύνδεσμος υψηλής ταχύτητας για την ανταλλαγή μηνυμάτων
  - Ο σύνδεσμος μπορεί να είναι ένα τοπικό δίκτυο
  - Κάθε υπολογιστής είναι πολυεπεξεργαστής (δεν είναι απαραίτητο)

# Συστάδες - Διατάξεις



- Μία συστάδα με **κοινόχρηστο δίσκο** με ένα σύνδεσμο υψηλής ταχύτητας για την ανταλλαγή μηνυμάτων
  - Το υποσύστημα δίσκων είναι τεχνολογίας **RAID** για να έχουμε **υψηλή διαθεσιμότητα**. Εάν ένας δίσκος υποστεί βλάβη το σύστημα συνεχίζει κανονικά

# Συστάδες - Διατάξεις

---

## Κατηγοριοποίηση Συσταδοποίησης βάσει λειτουργιών

- **Παθητική Αναμονή:** Ένας υπολογιστής είναι υπεύθυνος για τη διαχείριση ολόκληρου του φορτίου επεξεργασίας. Ο 2<sup>ος</sup> υπολογιστής παραμένει ανενεργός και αναλαμβάνει σε περίπτωση βλάβης του κύριου υπολογιστή
- Ο κύριος υπολογιστής στέλνει περιοδικά beacon μηνύματα στο δευτερεύων. Εάν ο δευτερεύων δε λάβει μήνυμα τότε ενεργοποιείται

# Συστάδες - Διατάξεις

---

## Κατηγοριοποίηση Συσταδοποίησης βάσει λειτουργιών

- **Πλεονεκτήματα Παθητικής Αναμονής:**
    - Εύκολη Υλοποίηση
    - Αυξάνει τη Διαθεσιμότητα, αλλά δεν βελτιώνει την απόδοση.
  - **Μειονεκτήματα Παθητικής Αναμονής :**
    - Ανενεργός ο δευτερεύων υπολογιστής
    - Αν τα δύο συστήματα δεν μοιράζονται κοινούς δίσκους, τότε η εφεδρική μηχανή παρέχει μία λειτουργική εφεδρεία, αλλά δεν έχει πρόσβαση σε καμία από τις βάσεις δεδομένων
- ❖ *Η παθητική αναμονή δεν αναφέρεται ως συστάδα*

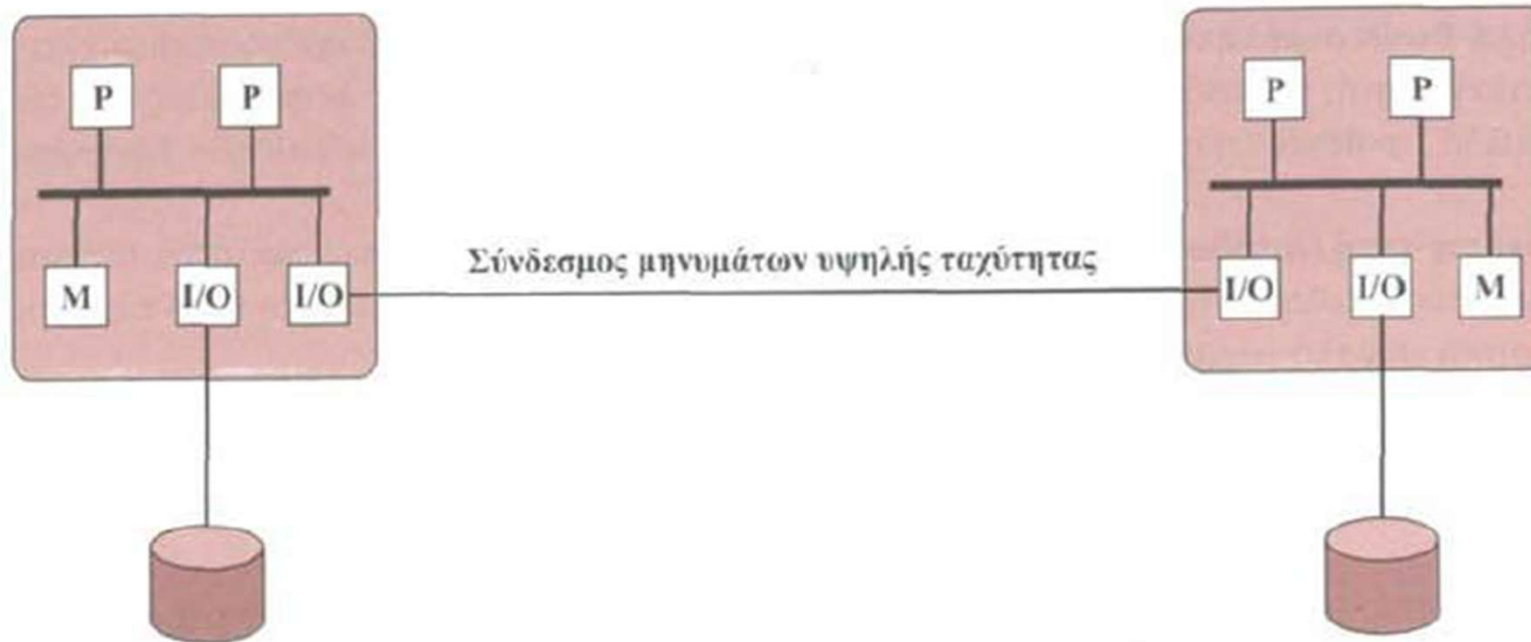
# Συστάδες - Διατάξεις

---

- Η συστάδα σχετίζεται με **ενεργούς** διασυνδεδεμένους υπολογιστές **που εκτελούν επεξεργασία δεδομένων**
- Η διάταξη αναφέρεται ως **Ενεργός Δευτερεύων Εξυπηρετητής**. Κατηγορίες Συσταδοποίησης:
  1. Ξεχωριστοί Εξυπηρετητές
  2. Τίποτε Κοινόχρηστο
  3. Κοινόχρηστη Μνήμη

# Συστάδες - Διατάξεις

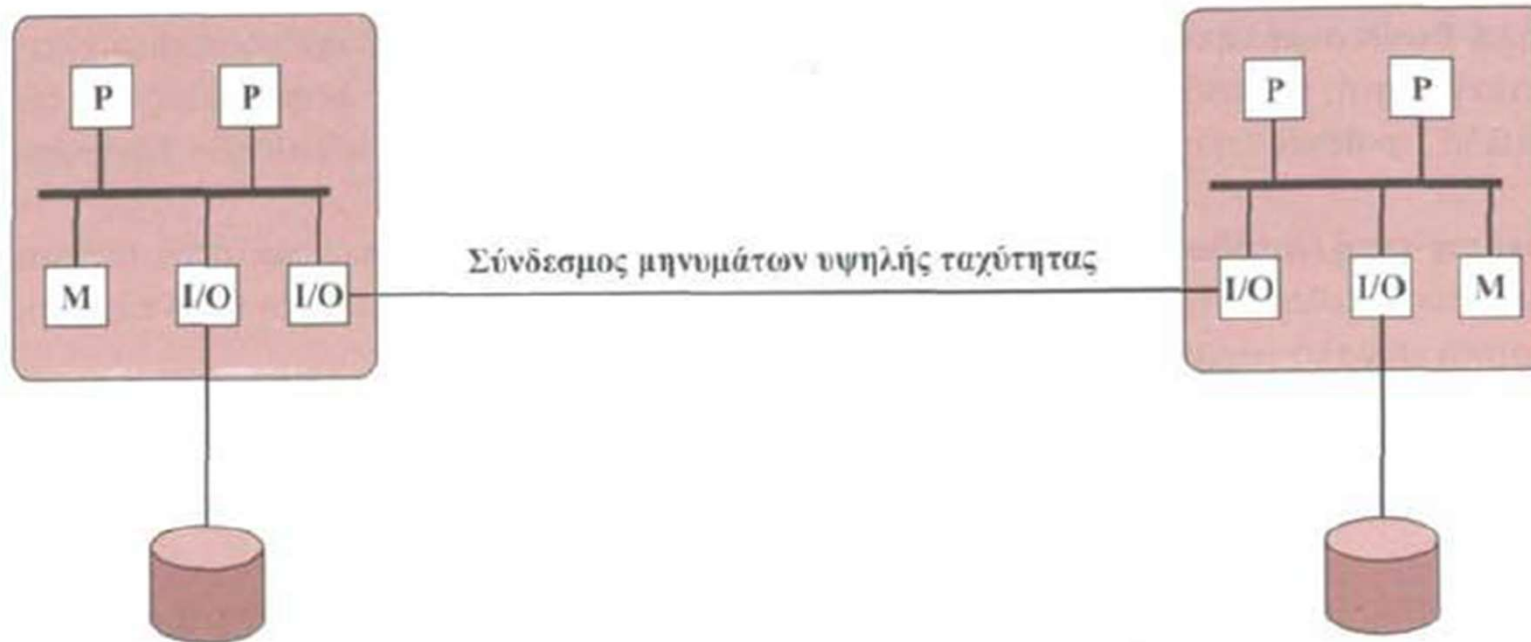
## Ξεχωριστοί Εξυπηρετητές (χωρίς κοινόχρηστους δίσκους)



- Τα δεδομένα πρέπει σε περιοδικά διαστήματα να αντιγράφονται από τον κύριο στο δευτερεύων Η/Υ ώστε αν πάθει βλάβη ένας Η/Υ να αναλάβει ένας άλλος
- + Υψηλή απόδοση και Διαθεσιμότητα

# Συστάδες - Διατάξεις

## Ξεχωριστοί Εξυπηρετητές (χωρίς κοινόχρηστους δίσκους)

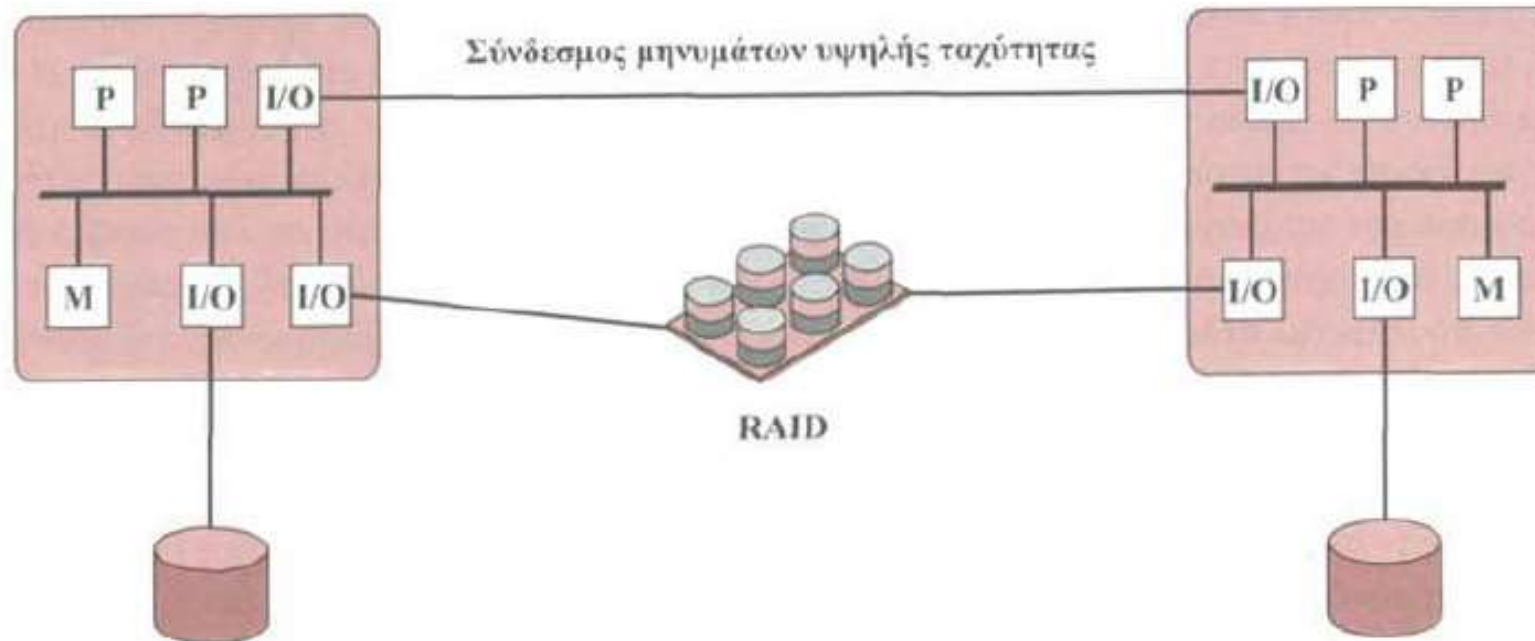


- Απαιτείται διαχείριση ή χρονοδρομολόγηση των εισερχόμενων αιτήσεων ώστε να έχουμε **load balancing**
- Μεγάλη επιβάρυνση στον εξυπηρετητή και στο δίκτυο λόγω ανταλλαγής δεδομένων



# Συστάδες - Διατάξεις

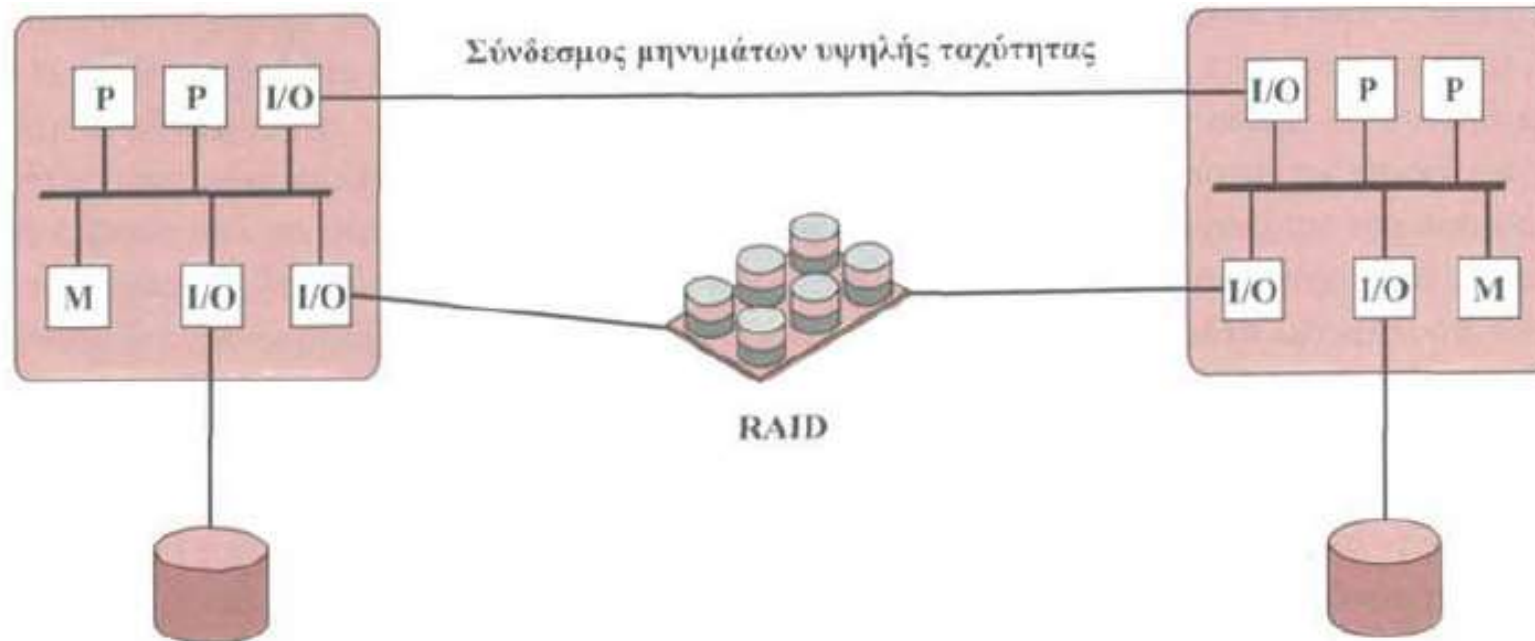
## Τίποτε Κοινόχρηστο (με κοινόχρηστους δίσκους)



- + Μείωση της επιβάρυνσης λόγω της επικοινωνίας καθώς οι εξυπηρετητές συνδέονται σε κοινόχρηστους δίσκους
- Οι **δίσκοι χωρίζονται σε τόμους**, και **κάθε τόμος ανήκει σε ένα Η/Υ**

# Συστάδες - Διατάξεις

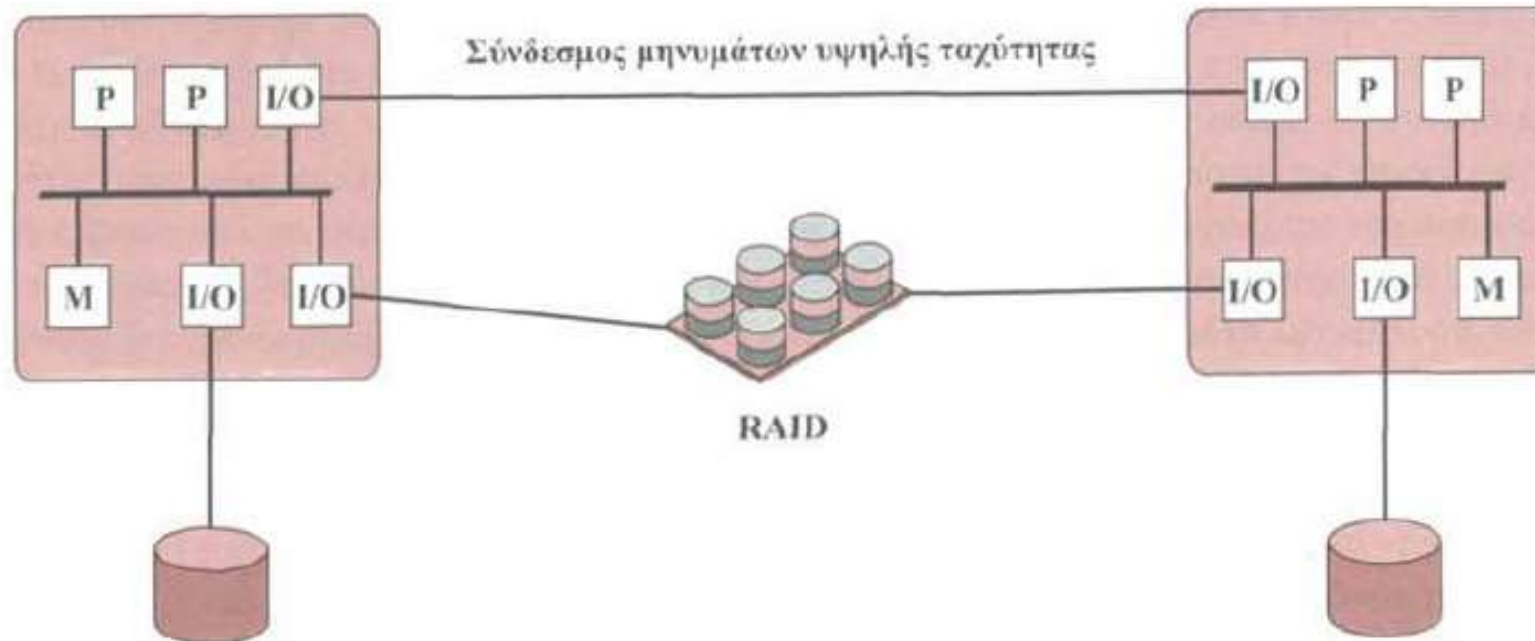
**Τίποτε Κοινόχρηστο** (με κοινόχρηστους δίσκους)



- Εάν ένας υπολογιστής υποστεί **βλάβη** τότε **ένας άλλος υπολογιστής λαμβάνει** στην κατοχή του **τους τόμους** εκείνου που έπαθε τη βλάβη

# Συστάδες - Διατάξεις

## Κοινόχρηστη Μνήμη (ονομάζεται και Κοινόχρηστοι Δίσκοι)



- Πολλοί εξυπηρετητές μοιράζονται ταυτόχρονα την προσπάθεια των δίσκων → Απαιτείται η δυνατότητα κλειδώματος των δεδομένων
- + Μειωμένη επιβάρυνση στο δίκτυο και στους εξυπηρετητές
- + Μειωμένος κίνδυνος διακοπής του συστήματος λόγω βλάβης ενός δίσκου

# Συστάδες και Λειτουργικό Σύστημα

---

## Διαχείριση Βλαβών

- Συστάδες υψηλής διαθεσιμότητας (1<sup>η</sup> προσέγγιση)
  - Μία συστάδα υψηλής διαθεσιμότητας *παρέχει μεγάλη πιθανότητα* όλοι οι *πόροι να είναι διαθέσιμοι*
  - Αν καταρρεύσει ένα σύστημα ή χαθεί ένας τόμος κάποιου δίσκου, τότε *χάνονται* όλα τα *ερωτήματα* το οποία *βρισκόταν σε επεξεργασία*
  - Αν τεθεί πάλι, *θα εξυπηρετηθεί από άλλο υπολογιστή*
  - Το λειτουργικό σύστημα της συστάδας *δεν παρέχει καμία εγγύηση για τις μερικώς εκτελεσμένες συναλλαγές*

# Συστάδες και Λειτουργικό Σύστημα

---

## Διαχείριση Βλαβών

- **Συστάδες με ανοχή στα σφάλματα (2<sup>η</sup> προσέγγιση)**
  - *Εξασφαλίζει* ότι όλοι οι *πόροι είναι πάντοτε διαθέσιμοι*
  - Επιτυγχάνεται με τη χρήση επιπλέον κοινόχρηστων δίσκων και μηχανισμών εφεδρικής αποθήκευσης τόσο για τις μη ολοκληρωμένες συναλλαγές όσο και για τις ολοκληρωμένες
  - **Βλάβη με μεταφορά:** Η λειτουργία της μεταφοράς εφαρμογών και πόρων δεδομένων από ένα σύστημα το οποίο έχει υποστεί ζημιά σε ένα εναλλακτικό σύστημα
  - **Βλάβη με επιστροφή:** Η αποκατάσταση των εφαρμογών και των δεδομένων στο αρχικό σύστημα αφού επιδιορθωθεί

# Συστάδες και Λειτουργικό Σύστημα

---

## Ισορροπία Φόρτου

- Απαιτείται να υπάρχει *load balancing* μεταξύ των υπολογιστών μίας συστάδας
- Όταν *προστίθεται* στη συστάδα ένας *νέος υπολογιστής*, θα πρέπει να *συμπεριλαμβάνεται* σε αυτή την *απαίτηση εξισορρόπησης*
- Είναι δυνατή η εμφάνιση υπηρεσιών σε διαφορετικά μέλη της συστάδας, αλλά και η μετακόμισή τους από ένα μέλος σε ένα άλλο
  - Το λογισμικό θα επιλαμβάνεται αυτών των θεμάτων

# Συστάδες και Λειτουργικό Σύστημα

---

## Παραλληλισμός Υπολογισμών

- Η αποτελεσματική χρήση μιας συστάδας απαιτεί σε ορισμένες περιπτώσεις την παράλληλη εκτέλεση του λογισμικού μιας εφαρμογής
- Τρεις προσεγγίσεις:
  1. Παραλληλοποίηση του Μεταγλωττιστή
  2. Παραλληλοποίηση της Εφαρμογής
  3. Παραμετρικοί Υπολογισμοί

# Συστάδες και Λειτουργικό Σύστημα

---

## Παραλληλοποίηση του Μεταγλωττιστή

- Ένας μεταγλωττιστής *προσδιορίζει*, κατά τη διάρκεια της μεταγλώττισης, *τα μέρη της εφαρμογής* που είναι δυνατόν *να εκτελεστούν παράλληλα*.
- Τα μέρη της εφαρμογής *εκχωρούνται σε ξεχωριστούς υπολογιστές της συστάδας*
- Η απόδοση εξαρτάται από τη φύση του προβλήματος και από το πόσο καλά είναι σχεδιασμένος ο μεταγλωττιστής
  - Οι μεταγλωττιστές αυτής της μορφής ***δεν είναι εύκολοι στην υλοποίηση***



# Συστάδες και Λειτουργικό Σύστημα

---

## Παραλληλοποίηση της Εφαρμογής

- Ο προγραμματιστής *γράφει την εφαρμογή* από την αρχή, έτσι *ώστε αυτή να εκτελείται σε μία συστάδα* και χρησιμοποιεί την *ανταλλαγή μηνυμάτων για τη μετακίνηση των δεδομένων* μεταξύ των κόμβων
- Επιβάρυνση του προγραμματιστή
- + Καλύτερη εκμετάλλευση της συστάδας

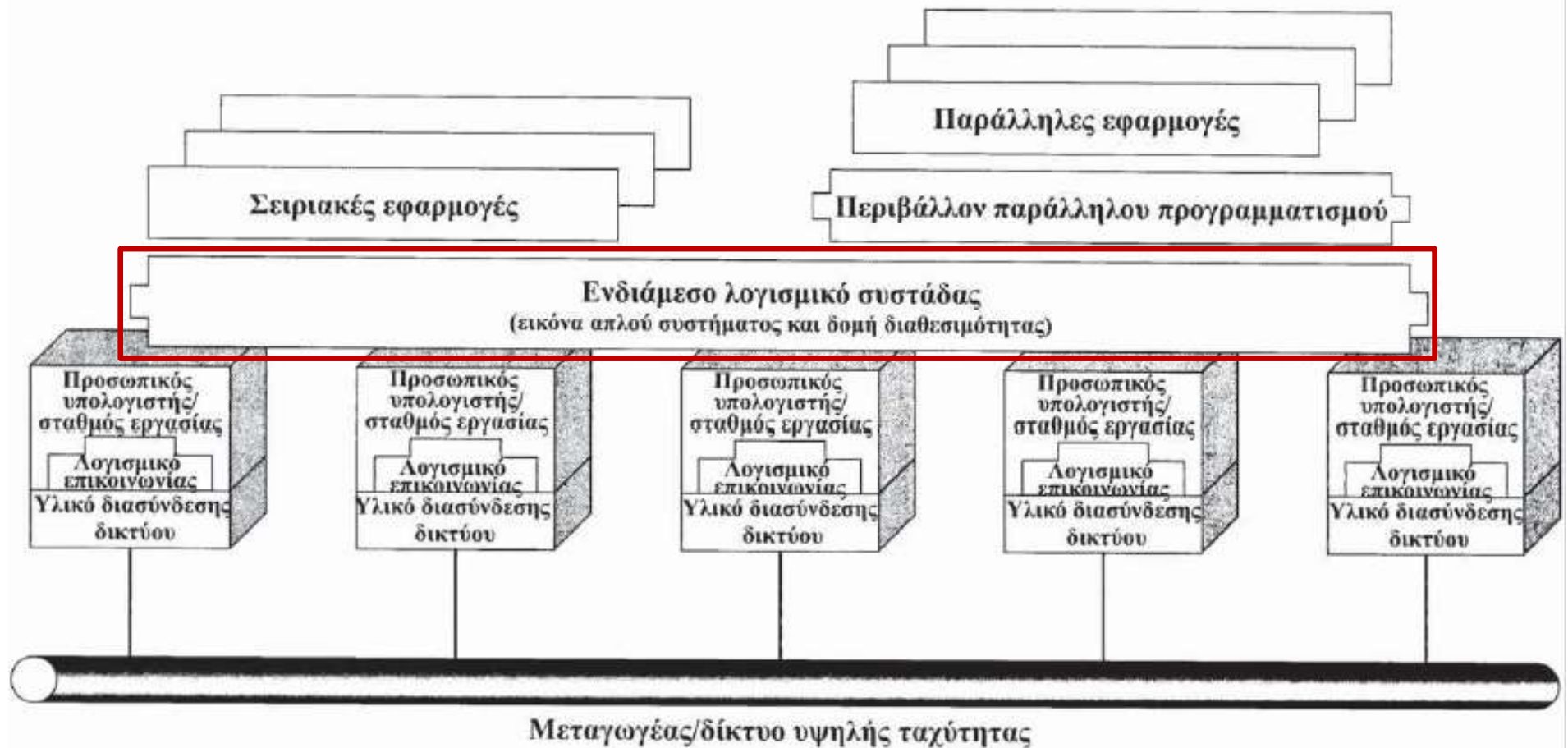
# Συστάδες και Λειτουργικό Σύστημα

---

## Παραμετρικοί Υπολογισμοί

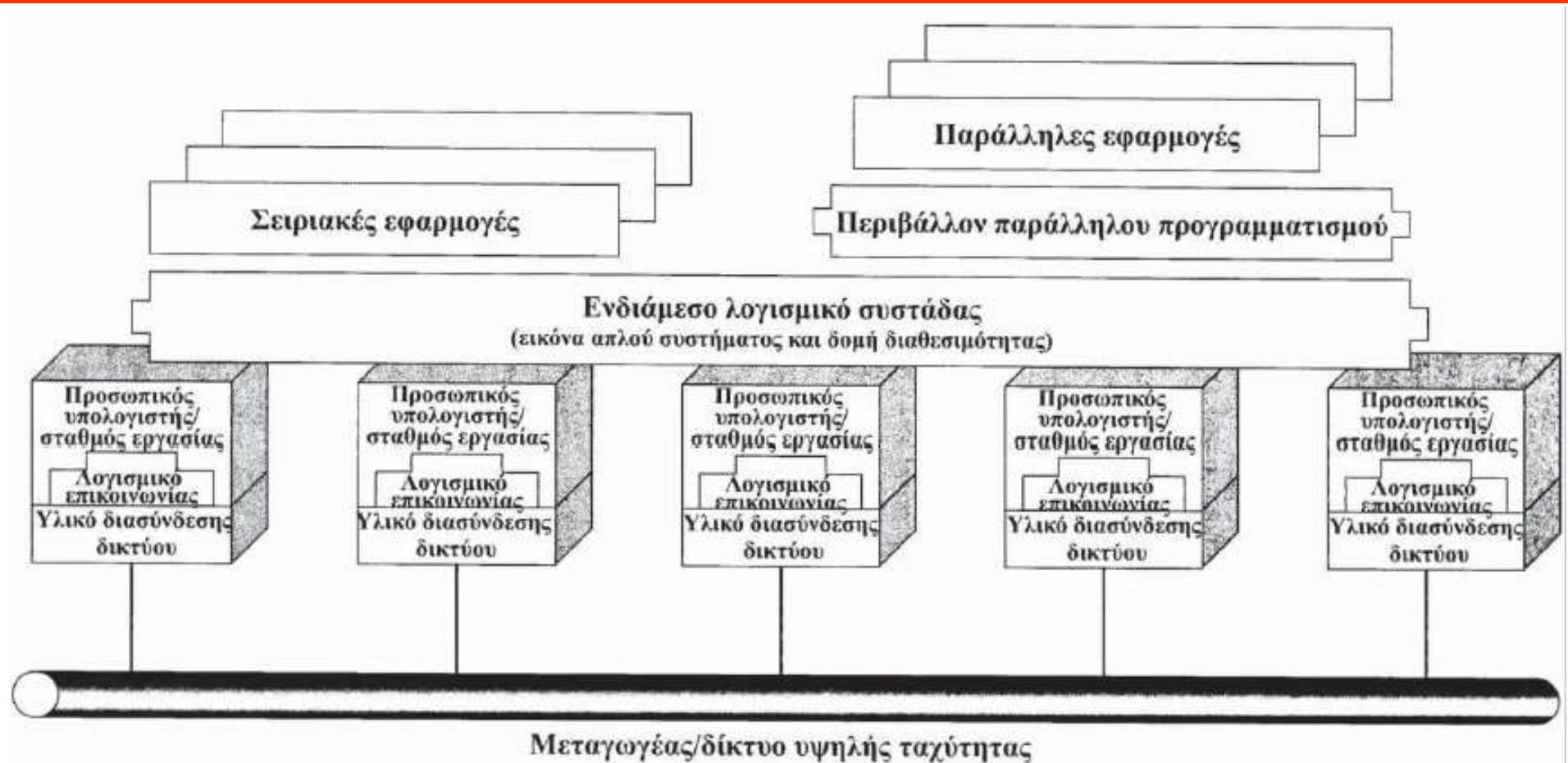
- Χρησιμοποιείται αν η ουσία της εφαρμογής είναι ένας *αλγόριθμος ή πρόγραμμα το οποίο πρέπει να εκτελεστεί πολλές φορές, κάθε φορά με διαφορετικό σύνολο συνθηκών και παραμέτρων εκκίνησης*
- Παράδειγμα: Ένα μοντέλο προσομοίωσης, το οποίο θα εκτελέσει ένα μεγάλο πλήθος διαφορετικών σεναρίων και θα δώσει συνοπτικά στατιστικά αποτελέσματα.

# Αρχιτεκτονική Συστάδας Υπολογιστών



- Κάθε υπολογιστής έχει εγκατεστημένο ένα **ενδιάμεσο στρώμα λογισμικού**, ώστε να **επιτρέπει τη λειτουργία της συστάδας**
- Το ενδιάμεσο αυτό στρώμα **παρέχει μία εικόνα ενοποιημένου συστήματος**

# Αρχιτεκτονική Συστάδας Υπολογιστών



- Το **ενδιάμεσο στρώμα** είναι επίσης υπεύθυνο για να **παρέχει υψηλή διαθεσιμότητα**, μέσα από *load balancing* και την **απόκριση σε βλάβες** των διαφόρων στοιχείων

# Αρχιτεκτονική Συστάδας Υπολογιστών

---

**Ενδιάμεσο Στρώμα Λογισμικού** – Επιθυμητές Υπηρεσίες & Λειτουργίες

- **Μονό σημείο εισόδου:** Ένας χρήστης εισέρχεται σε μία συστάδα και όχι σε μία ξεχωριστή μηχανή
- **Μονή ιεραρχία αρχείων:** Ο χρήστης βλέπει μία ιεραρχία καταλόγων αρχείων κάτω από τον ίδιο κατάλογο-ρίζα.
- **Μονό σημείο ελέγχου:** Υπάρχει ένας προεπιλεγμένος σταθμός εργασίας ο οποίος χρησιμοποιείται για τη διαχείριση και τον έλεγχο της συστάδας.

# Αρχιτεκτονική Συστάδας Υπολογιστών

---

**Ενδιάμεσο Στρώμα Λογισμικού** – Επιθυμητές Υπηρεσίες & Λειτουργίες

- **Μονή ιδεατή δικτύωση:** Κάθε κόμβος έχει τη δυνατότητα προσπέλασης κάθε άλλου σημείου της συστάδας. Υπάρχει μία απλή λειτουργία ιδεατού δικτύου.
- **Μονός χώρος μνήμης:** Η κατανεμημένη κοινόχρηστη μνήμη επιτρέπει στα προγράμματα να μοιράζονται μεταβλητές
- **Μονό σύστημα διαχείρισης εργασιών:** Μέσω ενός χρονοδρομολογητή εργασιών της συστάδας, ο χρήστης μπορεί να υποβάλλει μία εργασία χωρίς να καθορίσει τον υπολογιστή ο οποίος θα την εκτελέσει.

# Αρχιτεκτονική Συστάδας Υπολογιστών

---

**Ενδιάμεσο Στρώμα Λογισμικού** – Επιθυμητές Υπηρεσίες & Λειτουργίες

- **Μονή διεπαφή χρήστη:** Μία **μονή διεπαφή** με χρήση γραφικών **για όλους τους χρήστες**, ανεξάρτητα από το σταθμό από τον οποίο εισήλθαν στη συστάδα
- **Μονός χώρος E/E:** Κάθε κόμβος μπορεί να προσπελάσει **απομακρυσμένα** κάθε περιφερειακή μονάδα E/E ή κάθε δίσκο
  - χωρίς να γνωρίζει τη φυσική τους θέση
- **Μετακόμιση διεργασίας:** Αυτή η λειτουργία **επιτρέπει την εξισορρόπηση του φορτίου**

# Αρχιτεκτονική Συστάδας Υπολογιστών

---

**Ενδιάμεσο Στρώμα Λογισμικού** – Επιθυμητές Υπηρεσίες & Λειτουργίες

- **Μονός χώρος διεργασιών:** Μία διεργασία σε οποιονδήποτε κόμβο έχει τη δυνατότητα να δημιουργήσει μία άλλη διεργασία ή να επικοινωνήσει με μία άλλη διεργασία, σε έναν απομακρυσμένο κόμβο
  - Χρησιμοποιείται μία ομοιόμορφη τεχνική αναγνώρισης διεργασιών.
- **Έλεγχος σημείων:** Αυτή η λειτουργία αποθηκεύει περιοδικά την κατάσταση μίας διεργασίας και τα ενδιάμεσα αποτελέσματα των υπολογισμών, ώστε να υπάρχει δυνατότητα αποκατάστασης μετά από κάποια βλάβη.

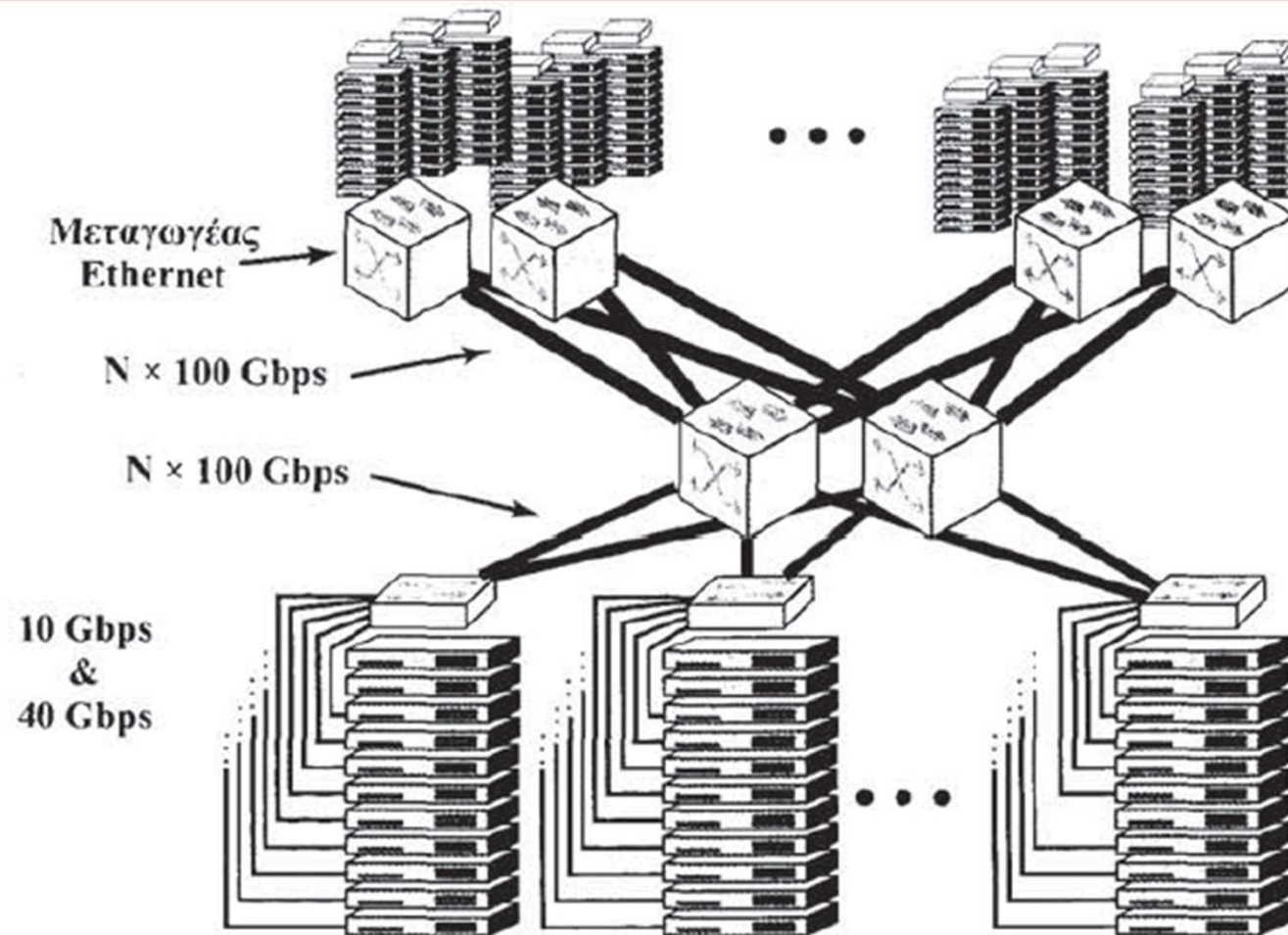


# Συστάδες – Φέτες Εξυπηρετητών

---

- Φέτες Εξυπηρετητών είναι μία υλοποίηση συστάδων
- Χρησιμοποιείται σε Data Center για εξοικονόμηση χώρου και καλύτερη διαχείριση του συστήματος
  - Πολλοί εξυπηρετητές («φέτες») σε ένα σασσί
  - Το σασσί παρέχει το ρεύμα
  - Κάθε φέτα έχει το δικό της εξυπηρετητή, μνήμη και δίσκο

# Συστάδες - Φέτες Εξυπηρετητών



- Θύρες **10Gbps** σε κάθε εξυπηρετητή
- **Μεταγωγείς Ethernet 100Gbps** χρησιμοποιούνται σε ανοδικούς συνδέσμους μέσα στο Data Center

# Σύγκριση Συστάδων και Συμμετρικών Πολυεπεξεργαστών

---

- Και οι δύο λύσεις παρέχουν μία **διάταξη πολλαπλών επεξεργαστών** για να υποστηρίξουν εφαρμογές υψηλών απαιτήσεων και είναι **διαθέσιμες στο εμπόριο**
- Ο συμμετρικός πολυεπεξεργαστής είναι πιο **εύκολος στη διαχείριση και στη διαμόρφωση**
- Ο συμμετρικός πολυεπεξεργαστής είναι **πιο κοντά στο μοντέλο του απλού επεξεργαστή** με βάσει το οποίο γράφονται συνήθως οι εφαρμογές
- Ο συμμετρικός πολυεπεξεργαστής **καταλαμβάνει πολύ λιγότερο χώρο** και **καταναλώνει λιγότερη ενέργεια**

# Σύγκριση Συστάδων και Συμμετρικών Πολυεπεξεργαστών

---

- **Εύκολη μετάβαση** από έναν απλό επεξεργαστή σε ένα **συμμετρικό πολυεπεξεργαστή**
  - Η βασική αλλαγή η οποία απαιτείται είναι η *λειτουργία του χρονοδρομολογητή*
- Τα προϊόντα **συμμετρικών πολυεπεξεργαστών** είναι **καθιερωμένα** και **σταθερά**
- Οι **συστάδες** είναι καλύτερες σχετικά με όρους **επαυξητικής και απόλυτης κλιμάκωσης**
- Οι **συστάδες** είναι καλύτερες με όρους **διαθεσιμότητας**

# Μη Ομοιόμορφη Προσπέλαση Μνήμης

---

- **Ομοιόμορφη προσπέλαση μνήμης:** Όλοι οι επεξεργαστές έχουν πρόσβαση σε όλα τα μέρη της κύριας μνήμης χρησιμοποιώντας λειτουργίες φόρτωσης και αποθήκευσης.
  - Ο χρόνος προσπέλασης κάθε περιοχής της μνήμης είναι ίδιος για κάθε επεξεργαστή
- **Μη ομοιόμορφη προσπέλαση μνήμης (NUMA):** Όλοι οι επεξεργαστές έχουν πρόσβαση σε όλα τα μέρη της κύριας μνήμης χρησιμοποιώντας λειτουργίες φόρτωσης και αποθήκευσης.
  - Ο χρόνος προσπέλασης της μνήμης διαφέρει αναλόγως με την περιοχή η οποία προσπελάυνεται

# Μη Ομοιόμορφη Προσπέλαση Μνήμης

---

- Μη ομοιόμορφη προσπέλαση μνήμης με συνοχή κρυφής μνήμης (CC-NUMA): Ένα σύστημα NUMA στο οποίο **διατηρείται η συνοχή ανάμεσα στις κρυφές μνήμες** των διαφόρων επεξεργαστών
  - Ένα σύστημα NUMA το οποίο δεν διαθέτει συνοχή κρυφής μνήμης, είναι περίπου παρόμοιο με μία συστάδα
- ❖ *Το ενδιαφέρον επικεντρώνεται στα συστήματα CC-NUMA*

# Μη Ομοιόμορφη Προσπέλαση Μνήμης

---

- Σε ένα σύστημα Συμμετρικών Πολυεπεξεργαστών μια αποτελεσματική τεχνική κρυφής μνήμης μειώνει την κυκλοφορία η οποία διεξάγεται μέσω του διαύλου, ανάμεσα σε κάποιον επεξεργαστή και την κύρια μνήμη
  - Όσο αυξάνεται το πλήθος των επεξεργαστών, αυξάνεται και η κυκλοφορία
- Επίσης, ο δίαυλος χρησιμοποιείται για την ανταλλαγή σημάτων τα οποία σχετίζονται με τη συνοχή της κρυφής μνήμης
- Ο δίαυλος μπορεί να αποτελέσει bottleneck → *Υπάρχει ένα όριο στο πλήθος των επεξεργαστών*

# Μη Ομοιόμορφη Προσπέλαση Μνήμης

---

- Το **όριο του πλήθους των επεξεργαστών** σε ένα σύστημα συμμετρικών πολυεπεξεργαστών **είναι ένα από τα κίνητρα τα οποία οδήγησαν στη δημιουργία των συστάδων**
- Στις **συστάδες** κάθε κόμβος διαθέτει τη δική του μνήμη και η **συνοχή διατηρείται σε επίπεδο λογισμικού και όχι υλικού**
- ❖ *Η διασπορά της μνήμης επηρεάζει την απόδοση*



# Μη Ομοιόμορφη Προσπέλαση Μνήμης

---

- Ο στόχος της προσέγγισης CC-NUMA είναι:
  1. να διατηρεί ένα διαφανές σύστημα μεγάλης μνήμης
  2. να επιτρέπει την ύπαρξη πολλαπλών κόμβων πολυεπεξεργαστών, καθένας από τους οποίους έχει το δικό του δίαυλο ή άλλο σύστημα διασύνδεσης
- Οργάνωση Συστήμα CC-NUMA:
  - πολλαπλοί ανεξάρτητοι κόμβοι,
  - κάθε κόμβος περιέχει πολλούς επεξεργαστές, καθένας εκ των οποίων διαθέτει τη δική του κρυφή μνήμη Επιπέδου 1 και 2
  - κάθε κόμβος περιέχει τη δική του κύρια μνήμη

# Μη Ομοιόμορφη Προσπέλαση Μνήμης

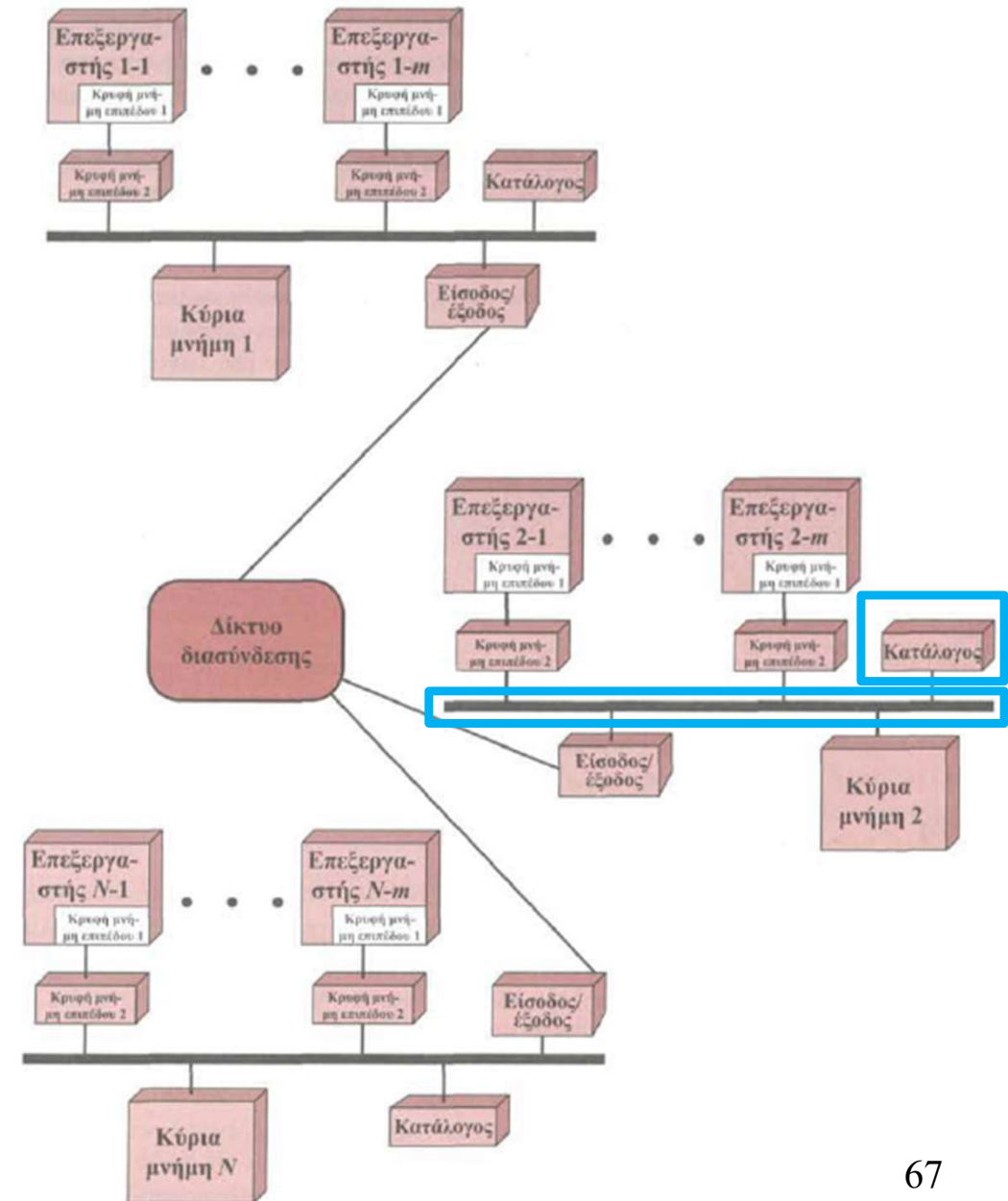
---

- Ένας επεξεργαστής έχει μία διευθυνσιοδοτούμενη μνήμη
  - η κάθε θέση της χαρακτηρίζεται από μία μοναδική διεύθυνση συστήματος.
- Ένας επεξεργαστής εκκινεί μία προσπέλαση της μνήμης:
  1. αν η θέση αυτή **δεν βρίσκεται στην κρυφή του μνήμη**, τότε η **κρυφή μνήμη Επιπέδου 2 ξεκινά μία λειτουργία ανάκλησης**
  2. Αν η επιθυμητή γραμμή βρίσκεται στο τοπικό κομμάτι της κύριας μνήμης, τότε **στέλνεται μία αυτόματη αίτηση στο δίκτυο διασύνδεσης**, ώστε να προσκομιστεί αυτή η γραμμή
  3. **να παραδοθεί στον τοπικό δίαυλο**, και στη συνέχεια, **να παραδοθεί στην κρυφή μνήμη** που έκανε την αίτηση

# Μη Ομοιόμορφη Προσπέλαση Μνήμης

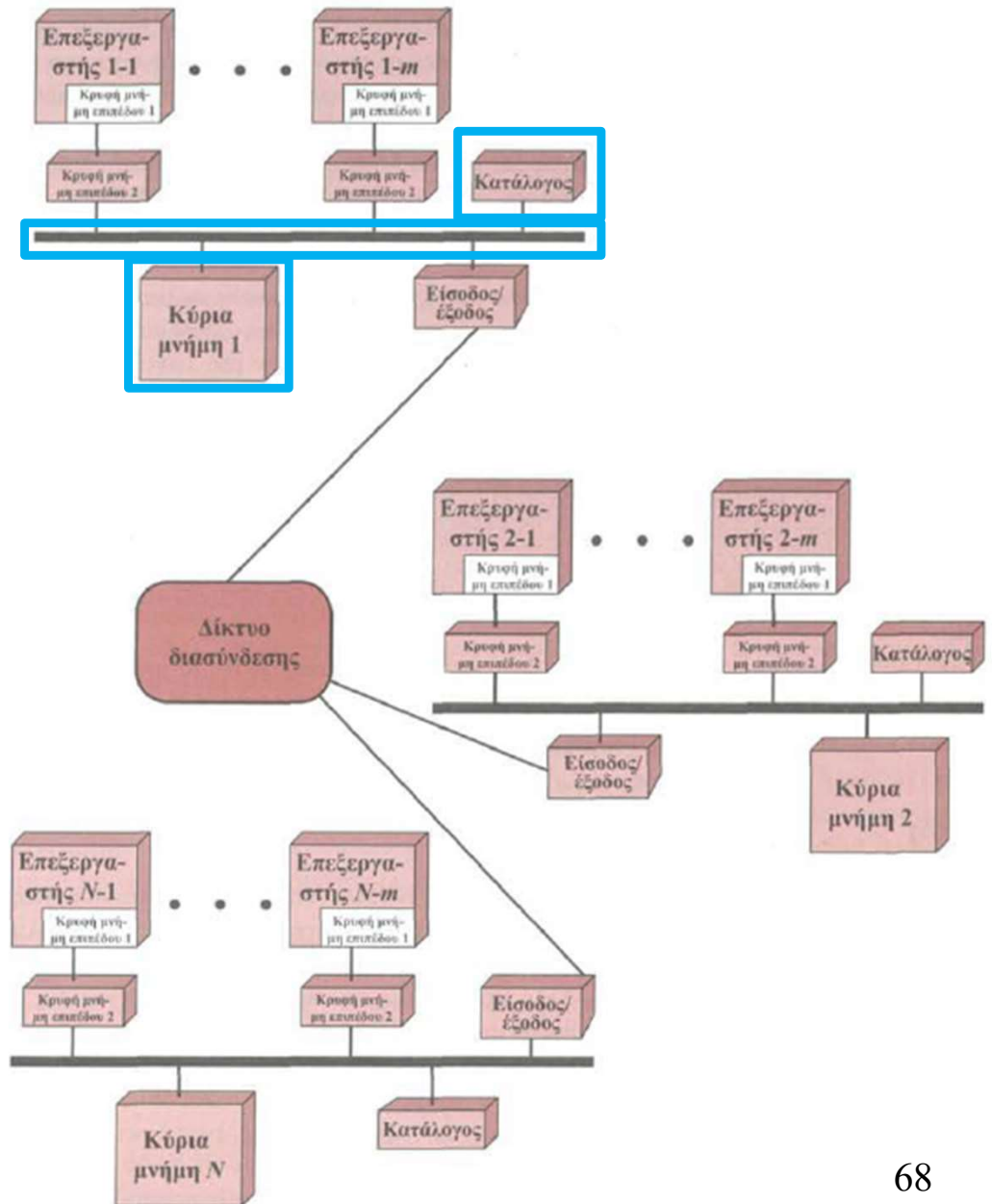
- Έστω Processor (P) 3 του κόμβου 2 (P2-3) ζητάει τη θέση μνήμης 798 που βρίσκεται στη μνήμη του κόμβου 1

1. Ο P2-3 εκδίδει μία αίτηση ανάγνωσης της θέσης 798, στο δίαυλο του κόμβου 2
2. Ο κατάλογος του κόμβου 2 βλέπει την αίτηση και αναγνωρίζει ότι η θέση βρίσκεται στον κόμβο 1



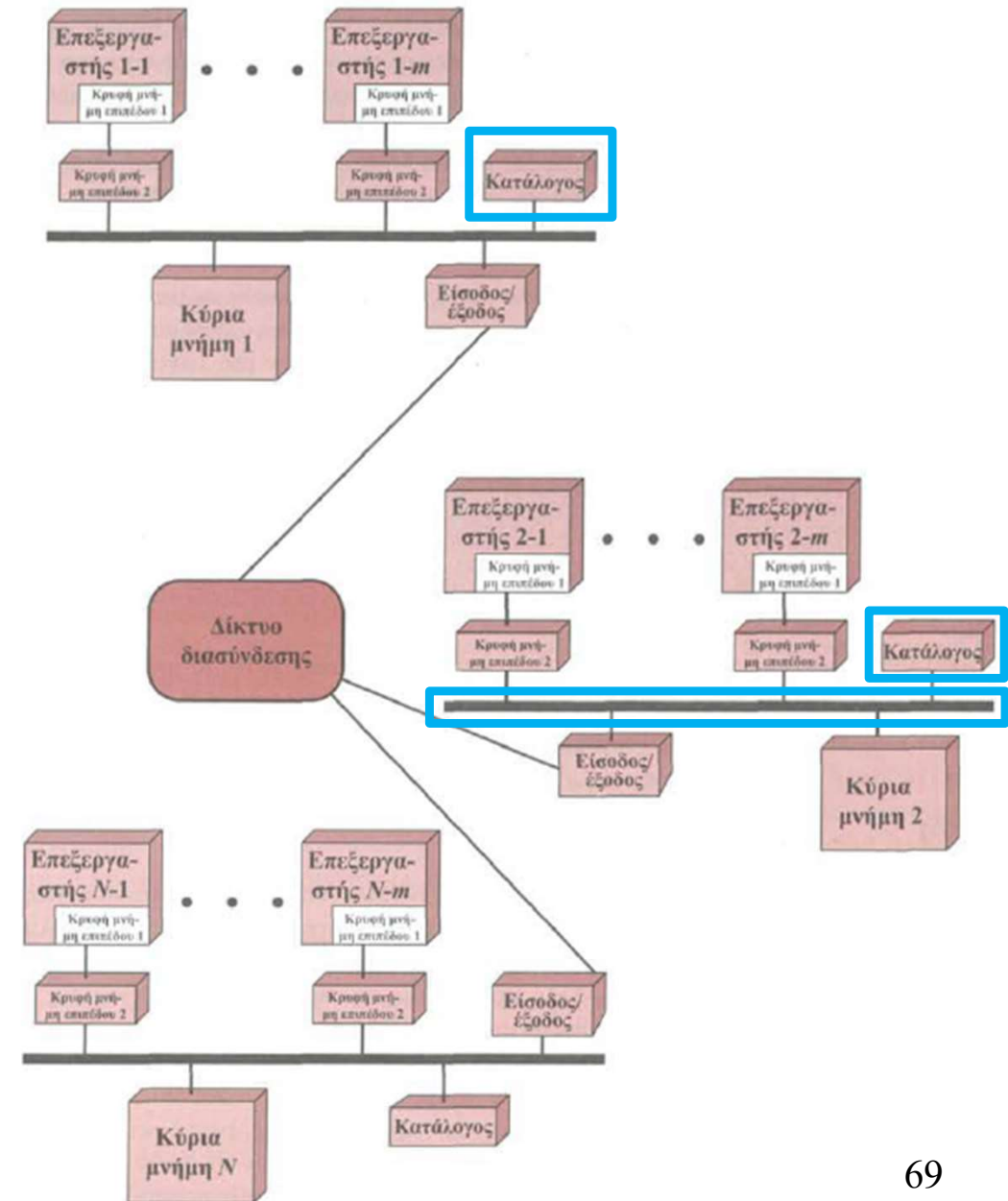
# Μη Ομοιόμορφη Προσπέλαση Μνήμης

3. Ο κατάλογος του κόμβου 2 στέλνει μία αίτηση στον κόμβο 1, η οποία λαμβάνεται από τον κατάλογο του κόμβου 1
4. Ο κατάλογος του κόμβου 1 ζητά τα περιεχόμενα της θέσης 798, σαν να ήταν ο ίδιος ο επεξεργαστής
5. Η κύρια μνήμη του κόμβου 1 αποκρίνεται τοποθετώντας τα δεδομένα που έχουν ζητηθεί στο δίαυλο.



# Μη Ομοιόμορφη Προσπέλαση Μνήμης

6. Ο κατάλογος του κόμβου 1 λαμβάνει τα δεδομένα από το δίαυλο
7. Η τιμή μεταφέρεται πίσω στον κατάλογο του κόμβου 2
8. Ο κατάλογος του κόμβου 2 τοποθετεί το μπλοκ δεδομένων στο δίαυλο του κόμβου 2, ενεργώντας ως πληρεξούσιος της μνήμης
9. Η τιμή λαμβάνεται και τοποθετείται στην κρυφή μνήμη του επεξεργαστή P2-3



# Μη Ομοιόμορφη Προσπέλαση Μνήμης

---

- Το παράδειγμα μας δείχνει **πως τα δεδομένα διαβάζονται από μία απομακρυσμένη μνήμη** χρησιμοποιώντας **μηχανισμούς του υλικού**, οι οποίοι καθιστούν τη συναλλαγή μη ορατή για τον επεξεργαστή
- Στην κορυφή αυτών των μηχανισμών, είναι αναγκαία η ύπαρξη μίας μορφής **πρωτοκόλλου συνοχής της κρυφής μνήμης**.
  - Διάφορα συστήματα διαφέρουν ακριβώς στο πρωτόκολλο συνοχής που χρησιμοποιούν

# NUMA Πλεονεκτήματα & Μειονεκτήματα

---

- + Έχει τη δυνατότητα να δίνει καλή απόδοση σε επίπεδα παραλληλισμού σε σχέση με το σύστημα συμμετρικών πολυεπεξεργαστών χωρίς να απαιτεί σημαντικές αλλαγές σε επίπεδο λογισμικού
- + Με πολλούς κόμβους NUMA η κυκλοφορία του διαύλου για κάθε κόμβο δε δημιουργεί πρόβλημα (περιορισμούς)
- Αν πολλές από τις προσπελάσεις μνήμης αφορούν απομακρυσμένους κόμβους, η απόδοση αρχίζει να μειώνεται

# Μη Ομοιόμορφη Προσπέλαση Μνήμης

---

Η μείωση είναι δυνατόν να αποφευχθεί

- Προσέγγιση 1:
  1. Η χρήση των κρυφών μνημών **Επιπέδου 1** και **2** ελαχιστοποιεί τις προσπελάσεις στη μνήμη, συμπεριλαμβανομένων και των απομακρυσμένων
  2. Αν ένα μεγάλο μέρος του λογισμικού διαθέτει **ικανοποιητική χρονική τοπικότητα**, τότε οι απομακρυσμένες προσπελάσεις μνήμης θα πρέπει να είναι λίγες



# Μη Ομοιόμορφη Προσπέλαση Μνήμης

---

Η μείωση είναι δυνατόν να αποφευχθεί

- Προσέγγιση 2:

1. αν το λογισμικό έχει ικανοποιητική **χωρική τοπικότητα** και αν χρησιμοποιείται η **ιδεατή μνήμη**, τότε **τα δεδομένα** τα οποία χρειάζονται σε μία εφαρμογή **θα βρίσκονται τοποθετημένα σε ένα περιορισμένο πλήθος πρόσφατα χρησιμοποιούμενων σελίδων**
2. Οι σελίδες είναι δυνατόν να **φορτωθούν** αρχικά **στην τοπική μνήμη της τρέχουσας εφαρμογής**