ΠΑΝΕΠΙΣΤΗΜΙΟ
ΔΥΤΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ

Συστήματα Παράλληλης και Κατανεμημένης Επεξεργασίας

Ενότητα: ΕΡΓΑΣΤΗΡΙΑΚΗ ΑΣΚΗΣΗ No:08

Δρ. Μηνάς Δασυγένης

mdasyg@ieee.org

Τμήμα Μηχανικών Πληροφορικής και Τηλεπικοινωνιών

Εργαστήριο Ψηφιακών Συστημάτων και Αρχιτεκτονικής Υπολογιστών

http://arch.icte.uowm.gr/mdasyg

Χρηματοδότηση

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

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Περιεχόμενα

1. Σκοπός της άσκησης

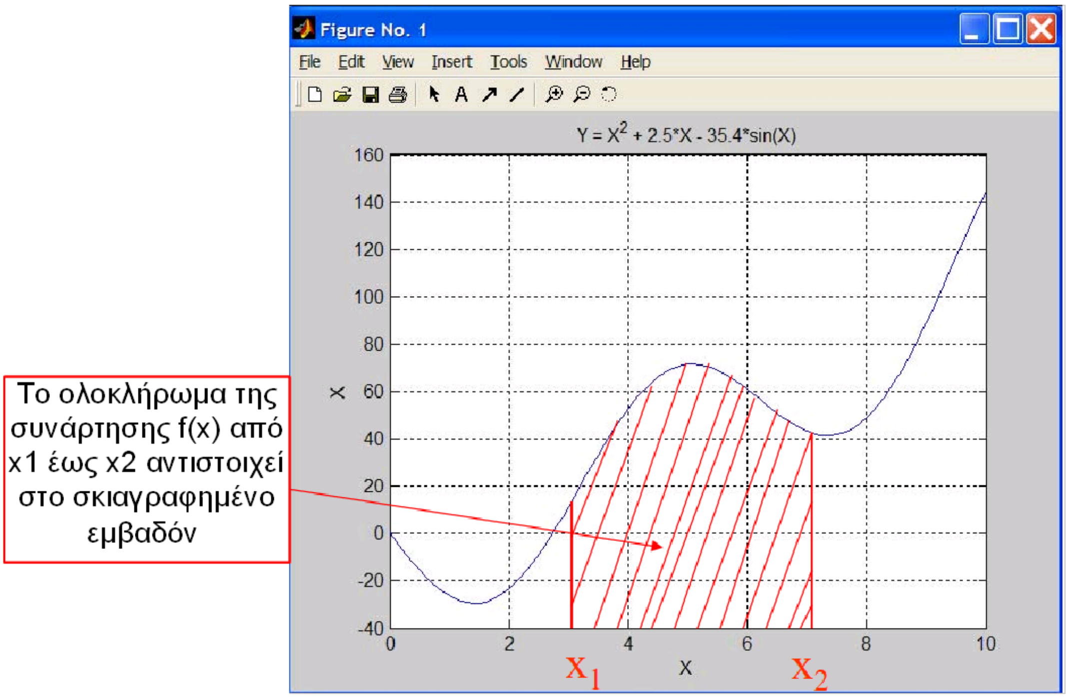
- Αριθμητική ολοκλήρωση με τη μέθοδο του τραπεζίου.
- Πολλαπλασιασμός πινάκων:
    - **(α)** με παρεμποδιστικές συναρτήσεις και
    - **(β)** συναρτήσεις συλλογικής επικοινωνίας.

2. Παραδοτέα

(A) 6 ερωτήσεις
(C) 5 ασκήσεις

3. Άσκηση 1: Ολοκλήρωση με τη μέθοδο του τραπεζίου

Να υπολογιστεί το ολοκλήρωμα $f(x) = \ 3x^2 + 1$ από 1 έως 4. Για να υπολογίσουμε το συγκεκριμένο ολοκλήρωμα θα χρησιμοποιήσουμε τη μέθοδο του τραπεζίου. Με τη μέθοδο του τραπεζίου μπορούμε κατά προσέγγιση να υπολογίσουμε ένα οποιοδήποτε ολοκλήρωμα πρώτου βαθμού. Στη μέθοδο αυτή υπολογίζουμε το εμβαδόν που βρίσκεται ανάμεσα στον άξονα x και στην καμπύλη f(x).



Για να υπολογίσουμε το συγκεκριμένο εμβαδόν χωρίζουμε την περιοχή σε πολλά υπο-τμήματα *(όσα περισσότερα τόσο καλύτερα)* μήκους **ΔΧ**.

Δημιουργούνται έτσι πολλά μικρά τραπέζια.

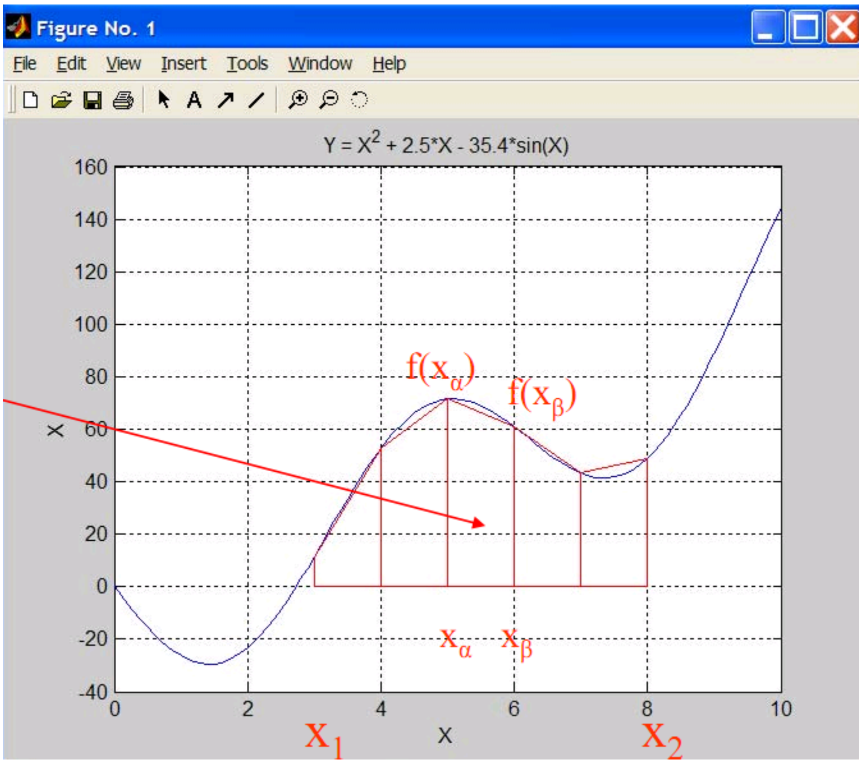
$$\int_{x1}^{x2} f(x) = \sum_{i=\tau\rho\alpha\pi\acute{\varepsilon}\zeta\iota o}^{\acute{o}\lambda\alpha\ \tau\alpha\ \tau\rho\alpha\pi\acute{\varepsilon}\zeta\iota\alpha} \boldsymbol{E\mu\beta\alpha\delta\acute{o}\ \tau\rho\alpha\pi\varepsilon\zeta\acute{\iota}o\upsilon\ i}$$

Το εμβαδό τραπεζίου υπολογίζεται από τον τύπο:

$$A_i = \frac{f\left(x_\alpha\right) + f\left(x_\beta\right)}{2} \times \Delta x$$

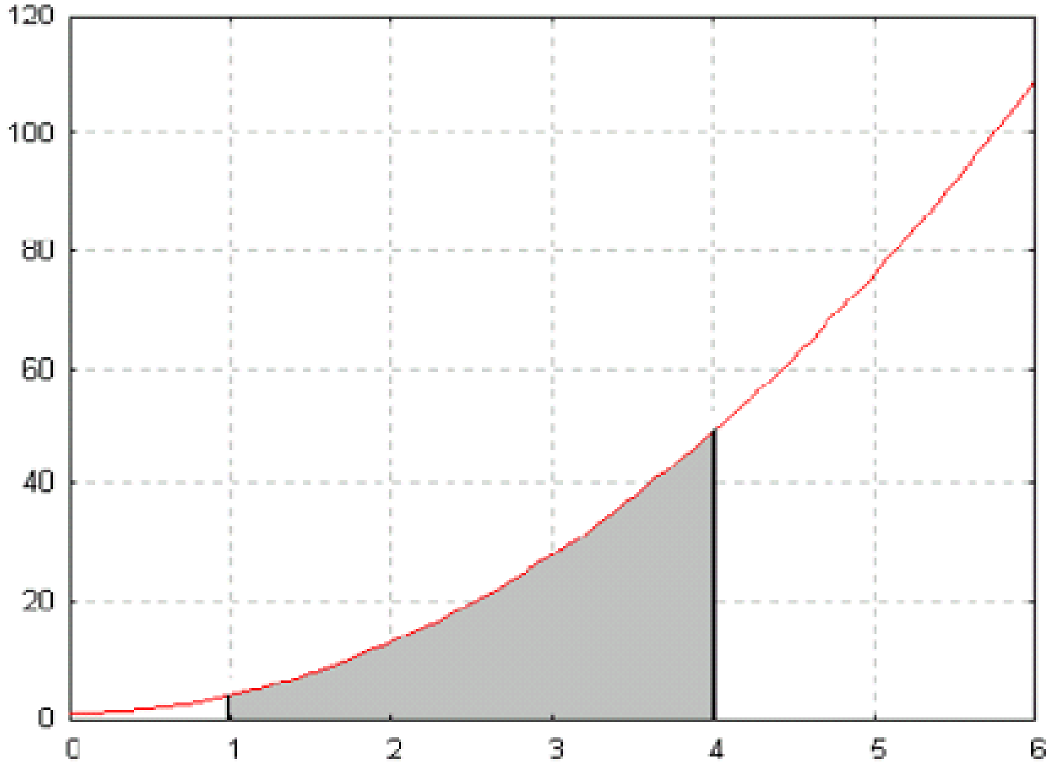
$$= \frac{f(x_\alpha) + f(x_\beta)}{2} \times \left(x_\beta - x_\alpha\right)$$

όπου $\mathbf{x_\alpha}$ και $\mathbf{x_\beta}$ αντιστοιχούν κάθε φορά στο τμήμα $[\mathbf{x_\alpha}, \mathbf{x_\beta}]$ που βρισκόμαστε, $\mathbf{\Delta x} = \mathbf{x_\beta} - \mathbf{x_\alpha}$ ενώ $\mathbf{f(x_\alpha)}$ και $\mathbf{f(x_\beta)}$ αντιστοιχούν στις τιμές που υπολογίζονται με τη χρήση της συνάρτησης.



Το ολοκλήρωμα λοιπόν της συνάρτησης ισούται με: $\int_{x1}^{x2} f(x) = \sum A_i$

Να υπολογιστεί το ολοκλήρωμα $f(x) = 3x^2 + 1$ από 1 έως 4



<u>Οδηγίες:</u>

- Θα πρέπει να υπολογιστεί το εμβαδόν που αντιστοιχεί στη γραμμοσκιασμένη περιοχή.
- **(C1)** Θα κατασκευάσετε ένα σειριακό πρόγραμμα *(χωρίς OpenMPI)* με όνομα **c1.c**. Το πρόγραμμα θα ρωτάει το χρήστη πόσα τραπέζια θα δημιουργήσει. (A1) Δοκιμάστε για 1, 10, 100, 1000, 10000 τραπέζια και σημειώστε τις τιμές.
- **(C2)** Αφού βεβαιωθείτε για την ορθή του λειτουργία, αντιγράψτε το προηγούμενο αρχείο στο αρχείο **c2.c** και τροποποιήστε το ώστε να εκτελεί παράλληλα τον υπολογισμό.
- **(A2)** Δοκιμάστε για 1, 10, 100, 1000, 10000 τραπέζια και σημειώστε τις τιμές και επιβεβαιώστε την ορθή λειτουργία του.
- **(A3)** Χρονομετρήστε τα 2 προγράμματα και βρείτε από ποιο αριθμό τραπεζίων και παραπάνω είναι πιο γρήγορο το παράλληλο πρόγραμμα.

*** Η τιμή που πρέπει να υπολογιστεί είναι 66.

4. Άσκηση 2: Πολλαπλασιασμός πινάκων

Θα αναπτυχθεί μια εφαρμογή που δέχεται ως είσοδο δυο δισδιάστατους πίνακες και θα υπολογίζει το γινόμενό τους. Για παράδειγμα, αν θεωρήσουμε για το σκοπό αυτό τους πίνακες:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots\cdots & a_{1N} \\ a_{21} & a_{22} & \cdots\cdots & a_{2N} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ a_{M1} & a_{M2} & \cdots\cdots & a_{MN} \end{pmatrix} \text{ και } B = \begin{pmatrix} b_{11} & b_{12} & \cdots\cdots & b_{1L} \\ b_{21} & b_{22} & \cdots\cdots & a_{2L} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ b_{K1} & b_{K2} & \cdots\cdots & a_{KL} \end{pmatrix}$$

με διαστάσεις **`MxN`** και **`KxL`** αντίστοιχα. Όπως είναι γνωστό από την άλγεβρα των πινάκων, ο πολλαπλασιασμός των πινάκων **`A`** και **`B`** είναι δυνατός μόνο όταν ο αριθμός των γραμμών του πίνακα **`B`** είναι ίσος με τον αριθμό των στηλών του πίνακα **`A`** – αυτό στην προκείμενη περίπτωση σημαίνει πως θα πρέπει να ισχύει η σχέση **`N=K`**. Στην περίπτωση κατά την οποία ισχύει αυτή η ισότητα, ο πίνακας που ορίζεται ως το γινόμενο των δύο πινάκων θα είναι ένας πίνακας C που θα έχει διαστάσεις **`MxL`** και στοιχεία:

$$C = \begin{pmatrix} c_{11} & c_{12} & \cdots\cdots & c_{1L} \\ c_{21} & c_{22} & \cdots\cdots & c_{2L} \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ c_{M1} & c_{M2} & \cdots\cdots & c_{ML} \end{pmatrix}$$

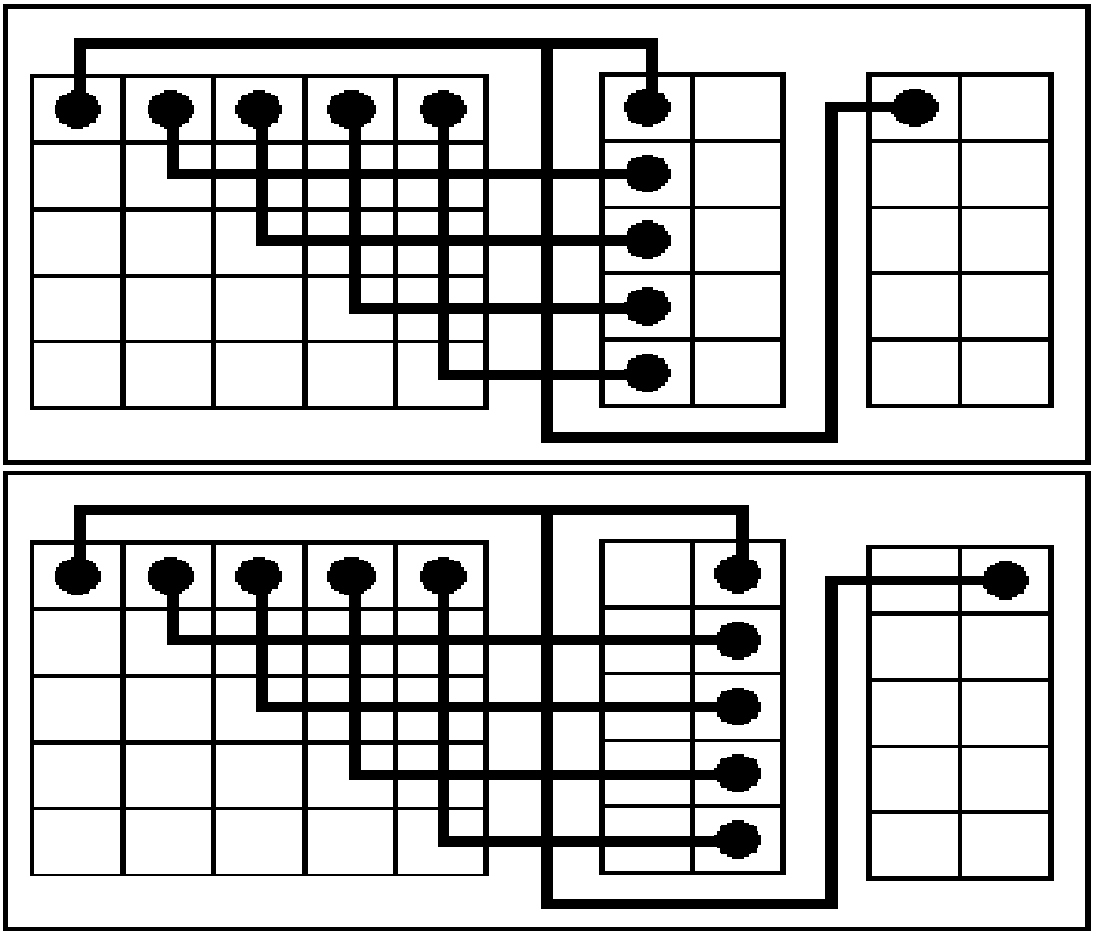
$$c_{ij} = \sum_{k=1}^{L} a_{ik} b_{kj} \text{ με i=1,2,…,L και j=1,2,3,…,M}$$

Δηλαδή, ο υπολογισμός του πολλαπλασιασμού αποτελείται από φωλιασμένους βρόχους 3 επιπέδων. Δηλαδή:

- για κάθε γραμμή του πίνακα A
    - για κάθε στήλη του πίνακα B
        - κάθε στήλη του A *(δηλαδή, κάθε στοιχείο της γραμμής)* => υπολόγισε το `c[i][j]`

Αφού υλοποιήσουμε τη σειριακή έκδοση θα ασχοληθούμε με την παράλληλη έκδοση. Για να το κάνουνε όμως αυτό, ας μελετήσουμε προσεκτικά τη διαδικασία υπολογισμού του γινομένου δύο πινάκων προκειμένου να διαπιστώσουμε ποια στάδια της διαδικασίας εκτελούνται ανεξάρτητα και χωρίς το ένα να επηρεάζει το άλλο, έτσι ώστε να μπορούν να πραγματοποιηθούν ταυτόχρονα από παράλληλες διεργασίες.

Η διαδικασία υπολογισμού του γινομένου δύο πινάκων παρουσιάζεται διαγραμματικά στο επόμενο σχήμα.



Στο σχήμα αυτό θεωρούμε την απλή περίπτωση δύο πινάκων εκ των οποίων ο πρώτος έχει διαστάσεις 5x5 ενώ ο δεύτερος έχει διαστάσεις 5x2 – επομένως το αποτέλεσμα του πολλαπλασιασμού αυτών των πινάκων θα είναι και αυτό ένας πίνακας με διαστάσεις 5x2. Από το παραπάνω σχήμα διαπιστώνουμε πως η διαδικασία υπολογισμού των στοιχείων της κάθε γραμμής του πίνακα γινομένου μπορεί να γίνει ανεξάρτητα από τον υπολογισμό των στοιχείων των υπολοίπων γραμμών *(το ίδιο βέβαια ισχύει και για τον υπολογισμό του κάθε στοιχείου της κάθε γραμμής ο οποίος μπορεί να γίνει ταυτόχρονα με τον υπολογισμό των υπολοίπων στοιχείων της ίδιας γραμμής του τελικού πίνακα – οι δύο ομάδες πράξεων που παρουσιάζονται στην άνω και στην κάτω εικόνα του ανωτέρω σχήματος μπορούν να γίνουν παράλληλα καθώς η μία δεν εξαρτάται από την άλλη)*. Σε πλήρη εφαρμογή λοιπόν των αλγορίθμων της παράλληλης επεξεργασίας, μπορούμε να κατασκευάσουμε ένα πλήθος παράλληλων διεργασιών οι οποίες θα αναλάβουν να υπολογίσουν τα στοιχεία των γραμμών του τελικού πίνακα που θα προκύψει από το γινόμενο των δύο πινάκων A και B.

Στην παράλληλη έκδοση λοιπόν μια διεργασία θα στέλνει σε κάθε άλλη διεργασία τη γραμμή του A πίνακα που πρέπει να υπολογίσει μαζί με όλον τον πίνακα B. Θα πρέπει να κοιτάξετε την περίπτωση που ο αριθμός των γραμμών του πίνακα A δε διαιρείται ακριβώς με τον αριθμό των διεργασιών που χρησιμοποιούμε. Σε αυτήν την περίπτωση κάποιες διεργασίες θα υπολογίσουν περισσότερα στοιχεία του τελικού πίνακα από τις άλλες. Θα πρέπει λοιπόν να υπολογίσουμε το υπόλοιπο της διαίρεσης αριθμός Γραμμών πίνακα A δια αριθμό διεργασιών. Αν το υπόλοιπο είναι 0 τότε η κάθε διεργασία θα έχει ίσο αριθμό γραμμών. Αν είναι διαφορετική από το 0 έστω K, τότε οι K πρώτες διεργασίες θα πρέπει να υπολογίσουν ένα στοιχείο παραπάνω.

Επίσης, θα πρέπει να θυμηθούμε ότι οι πίνακες **`MxN`** αποθηκεύονται σε μια συνεχόμενη περιοχή μνήμης **`M*N`** ως μια πολύ μεγάλη μονοδιάστατη περιοχή μνήμης.

Θα πρέπει να χρησιμοποιήσετε την παράμετρο **`tag`**.

Οδηγίες:

- **(C3)** Κατασκευάσατε τη σειριακή έκδοση του προγράμματος **c3.c**. Το πρόγραμμα θα διαβάσει από δυο αρχεία `arrayA.dat`, `arrayB.dat` τις πραγματικές τιμές (float) των δυο πινάκων. Τα αρχεία αυτά θα έχουν γραμμές και στήλες όπως θα απεικονίζαμε τους πίνακες. Θα χρησιμοποιεί **`#define`** για να ορίζει τις γραμμές και στήλες των πινάκων A και B.
  (A4) Μετρήστε το χρόνο εκτέλεσης χρησιμοποιώντας τις κατάλληλες συναρτήσεις.

- **(C4)** Αφού βεβαιωθείτε για την ορθή λειτουργία αντιγράψτε το παραπάνω αρχείο στο αρχείο **c4.c**. Η έκδοση αυτή δε θα χρησιμοποιεί τις συλλογικές επικοινωνίες, αλλά τις παρεμποδιστικές (send/recv).
  (A5) Μετρήστε το χρόνο εκτέλεσης χρησιμοποιώντας τις κατάλληλες συναρτήσεις και επιβεβαιώστε την ορθή λειτουργία.

- **(C5)** Αντιγράψτε το παραπάνω αρχείο στο αρχείο **c5.c**. Στο πρόγραμμα αυτό θα χρησιμοποιούνται συναρτήσεις συλλογικής επικοινωνίας. Συγκεκριμένα θα χρησιμοποιείται η **`MPI_Bcast`** για να σταλεί ο πίνακας B και η **`MPI_Scatter`** για τη διασπορά του πίνακα A. Ο περιορισμός είναι ότι οι γραμμές του πίνακα A είναι ακέραιο πολλαπλάσιο του αριθμού των διεργασιών.
  (A6) Μετρήστε το χρόνο εκτέλεσης χρησιμοποιώντας τις κατάλληλες συναρτήσεις και επιβεβαιώστε την ορθή λειτουργία.