Parallel and Distributed Computing

Prodotto Matrice-Matrice

Prof. G. Laccetti

a.a. 2023-2024

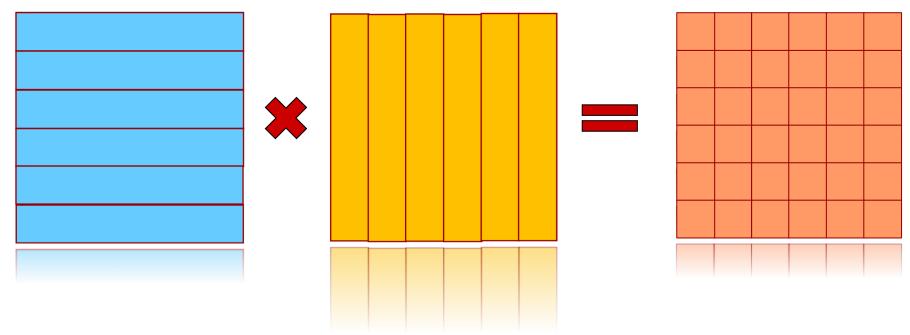
Problema

Progettazione
di un algoritmo parallelo
per architettura MIMD
a memoria distribuita
per il calcolo del prodotto righe per colonne
di 2 matrici A e B:

$$C = A \bullet B$$
, $A, B \in \mathbb{R}^{n \times n}$

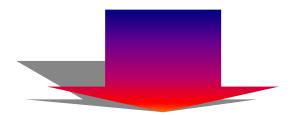
Quali sono i sotto-problemi indipendenti?

Gli elementi di C sono
calcolati effettuando i
prodotti scalari di ciascuna riga di A per
ciascuna colonna di B



Quali sono i sotto-problemi indipendenti?

Gli elementi di C sono
calcolati effettuando i
prodotti scalari di ciascuna riga di A per
ciascuna colonna di B



I prodotti scalari sono calcolati

in maniera indipendente

l'uno dall'altro

Altra possibile decomposizione

Decomposizione in BLOCCHI di RIGHE



Decomposizione in BLOCCHI di COLONNE



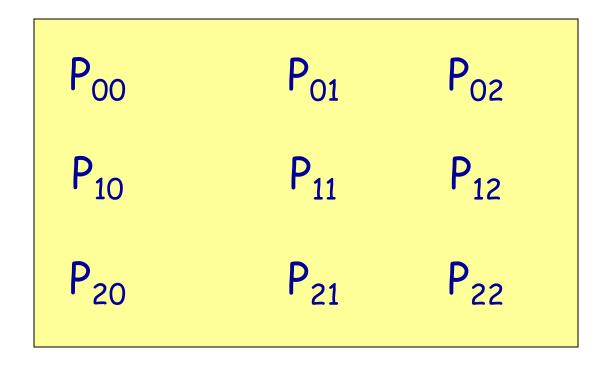
Decomposizione in BLOCCHI QUADRATI

Esempio: matrici a blocchi 3x3 = 9

$$\begin{bmatrix} \boldsymbol{C}_{00} & \boldsymbol{C}_{01} & \boldsymbol{C}_{02} \\ \boldsymbol{C}_{10} & \boldsymbol{C}_{11} & \boldsymbol{C}_{12} \\ \boldsymbol{C}_{20} & \boldsymbol{C}_{21} & \boldsymbol{C}_{22} \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_{00} & \boldsymbol{A}_{01} & \boldsymbol{A}_{02} \\ \boldsymbol{A}_{10} & \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{20} & \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{B}_{00} & \boldsymbol{B}_{01} & \boldsymbol{B}_{02} \\ \boldsymbol{B}_{10} & \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{B}_{20} & \boldsymbol{B}_{21} & \boldsymbol{B}_{22} \end{bmatrix}$$

Supponiamo di disporre di pxp processori secondo una griglia di dimensione pxp...

Esempio: 3x3=9 processori



Al processore P_{ij} assegniamo i blocchi A_{ij} e B_{ij}

Distribuzione dei dati

P _{oo}	P _{O1}	P ₀₂
$A_{00} B_{00}$	$A_{01} B_{01}$	$A_{02} B_{02}$
P ₁₀	P ₁₁	P ₁₂
A ₁₀ B ₁₀	A_{11} B_{11}	A_{12} B_{12}
P ₂₀	P ₂₁	P ₂₂
A ₂₀ B ₂₀	A ₂₁ B ₂₁	A ₂₂ B ₂₂

IDEA!

Con i dati così distribuiti $\begin{tabular}{ll} Vogliamo che il processore P_{ij} \\ Calcoli il blocco C_{ij} \\ \end{tabular}$

Ovvero...

$$C_{00} = A_{00} B_{00} + A_{01} B_{10} + A_{02} B_{20}$$

$$C_{01} = C_{02} =$$
 $A_{00}B_{01} + A_{01}B_{11} + A_{02}B_{21}$
 $A_{00}B_{02} + A_{01}B_{12} + A_{02}B_{22}$

$$C_{02} = A_{00}B_{02} + A_{01}B_{12} + A_{02}B_{22}$$

$$C_{10} = A_{10} B_{00} + A_{11} B_{10} + A_{12} B_{20}$$

$$C_{11} = A_{10} B_{01} + A_{11} B_{11} + A_{12} B_{21}$$

$$C_{12} = A_{10} B_{02} + A_{11} B_{12} + A_{12} B_{22}$$

$$C_{20} = A_{20}B_{00} + A_{21}B_{10} + A_{22}B_{20}$$

$$C_{21} = A_{20}B_{01} + A_{21}B_{11} + A_{22}B_{21}$$

$$C_{22} = A_{20} B_{02} + A_{21} B_{12} + A_{22} B_{22}$$

Ovvero...

$$A_{00} B_{00}$$

$$C_{00} = A_{00} B_{00} + A_{01} B_{10} + A_{02} B_{20}$$

$$A_{01} B_{01}$$

$$C_{01} = A_{00}B_{01} + A_{01}B_{11} + A_{02}B_{21}$$

$$A_{02} B_{02}$$
 $C_{02} =$
 $A_{00}B_{02} + A_{01}B_{12} + A_{02}B_{22}$

$$A_{10} B_{10}$$
 $C_{10} =$
 $A_{10} B_{00} + A_{11} B_{10} + A_{12} B_{20}$

$$A_{11} B_{11}$$
 $C_{11} =$
 $A_{10} B_{01} + A_{11} B_{11} + A_{12} B_{21}$

$$A_{12} B_{12}$$

$$C_{12} =$$

$$A_{10} B_{02} + A_{11} B_{12} + A_{12} B_{22}$$

$$A_{20} B_{20}$$

$$C_{20} = A_{20}B_{00} + A_{21}B_{10} + A_{22}B_{20}$$

$$A_{21} B_{21}$$

$$C_{21} = A_{20}B_{01} + A_{21}B_{11} + A_{22}B_{21}$$

$$C_{22} = A_{20} B_{02} + A_{21} B_{12} + A_{22} B_{22}$$

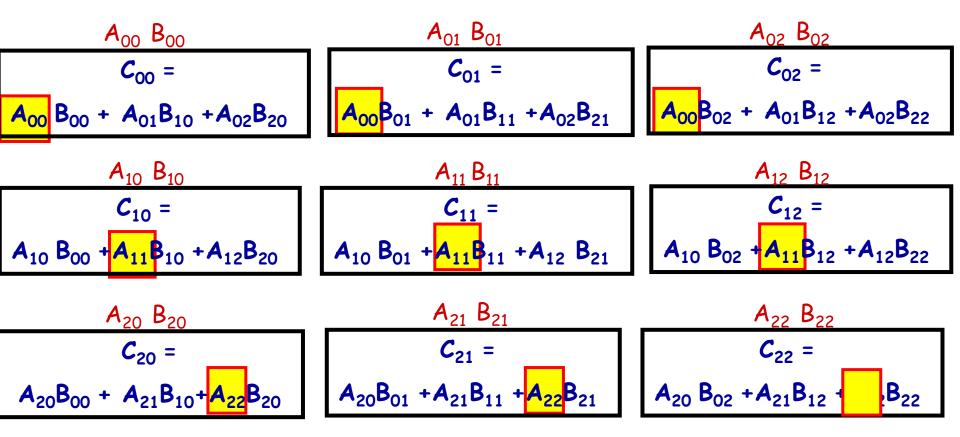
Osservazione I

Con la distribuzione dei dati effettuata solo i processori sulla diagonale della griglia, ovvero Pii possono calcolare "un contributo" del blocco ${\cal C}_{ii}$

Ovvero...

ovvero...

Osservazione II



Nel calcolo dei Cii ...

- $\cdot A_{00}$ è presente nel calcolo della I riga di C (C_{00}, C_{01}, C_{02})
- $\cdot A_{11}$ è presente nel calcolo della II riga di C (C_{10} , C_{11} , C_{12})
- · Agzècpresente nel calcolo della III riga di C (C20, C21, C22)

IDEA!

I processori sulla diagonale:

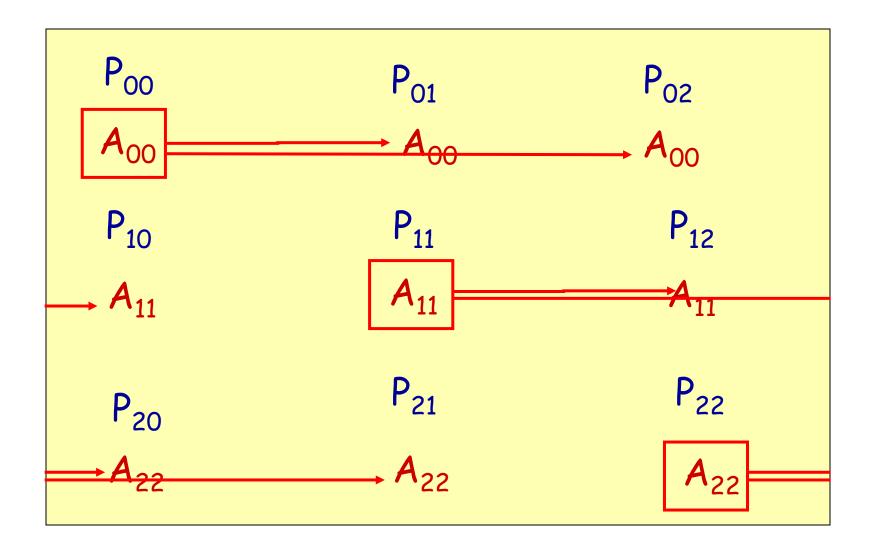
 P_{00} , P_{11} e P_{22}

inviano il proprio blocco di A,

ovvero A_{00} , $A_{11} e A_{22}$

Rispettivamente ai processori della riga 0,1,2 (ovvero ai processori che si trovano sulla loro stessa riga!)

I passo: invio dei blocchi diagonali di A



$$A_{00} B_{00}$$

$$C_{00} = A_{00}B_{00} + A_{01}B_{10} + A_{02}B_{20}$$

$$A_{00} A_{01} B_{01}$$
 $C_{01} =$
 $A_{00} B_{01} + A_{01} B_{11} + A_{02} B_{21}$

$$A_{00} A_{02} B_{02}$$

$$C_{02} =$$

$$A_{00} B_{02} + A_{01} B_{12} + A_{02} B_{22}$$

$$A_{11} A_{10} B_{10}$$

$$C_{10} =$$

$$A_{10} B_{00} + A_{11} B_{10} + A_{12} B_{20}$$

$$C_{11} = A_{10} B_{01} + A_{11} B_{11} + A_{12} B_{21}$$

$$C_{12} = A_{10} B_{02} + A_{11} B_{12} + A_{12} B_{22}$$

$$A_{22} A_{20} B_{20}$$

$$C_{20} =$$

$$A_{20} B_{00} + A_{21} B_{10} + A_{22} B_{20}$$

$$A_{22} A_{21} B_{21}$$

$$C_{21} = A_{20} B_{01} + A_{21} B_{11} + A_{22} B_{21}$$

$$A_{22} B_{22}$$

$$C_{22} = A_{20}B_{02} + A_{21}B_{12} + A_{22}B_{22}$$

$$A_{00} B_{00}$$

$$C_{00} = A_{00}B_{00} + A_{01}B_{10} + A_{02}B_{20}$$

$$A_{00} A_{01} B_{01}$$
 $C_{01} =$
 $A_{00} B_{01} + A_{01} B_{11} + A_{02} B_{21}$

$$A_{00} A_{02} B_{02}$$

$$C_{02} =$$

$$A_{00} B_{02} + A_{01} B_{12} + A_{02} B_{22}$$

$$A_{11} A_{10} B_{10}$$

$$C_{10} =$$

$$A_{10} B_{00} + A_{11} B_{10} + A_{12} B_{20}$$

$$A_{11}B_{11}$$
 $C_{11} = A_{10}B_{01} + A_{11}B_{11} + A_{12}B_{21}$

$$A_{11} A_{12} B_{12}$$

$$C_{12} =$$

$$A_{10} B_{02} + A_{11} B_{12} + A_{12} B_{22}$$

$$A_{22} A_{20} B_{20}$$

$$C_{20} =$$

$$A_{20} B_{00} + A_{21} B_{10} + A_{22} B_{20}$$

$$A_{22} A_{21} B_{21}$$

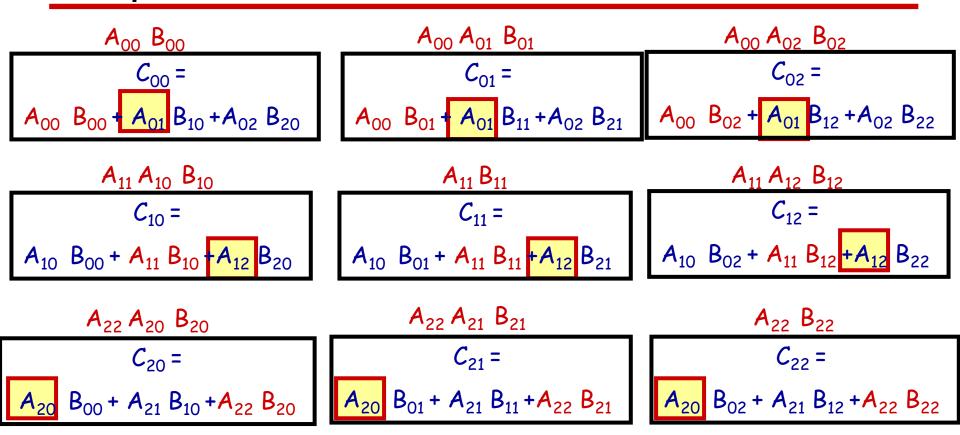
$$C_{21} = A_{20} B_{01} + A_{21} B_{11} + A_{22} B_{21}$$

$$A_{22} B_{22}$$

$$C_{22} =$$

$$A_{20}B_{02} + A_{21}B_{12} + A_{22}B_{22}$$

II passo



Nel calcolo dei C_{ij} ...

- • A_{01} è presente nel calcolo della I riga di C (C_{00}, C_{01}, C_{02})
 • A_{12} è presente nel calcolo della II riga di C (C_{10}, C_{11}, C_{12})
- G. LACCETTE presente nel calcolo della III riga di C (C20, C21, C22)

IDEA!

I processori:

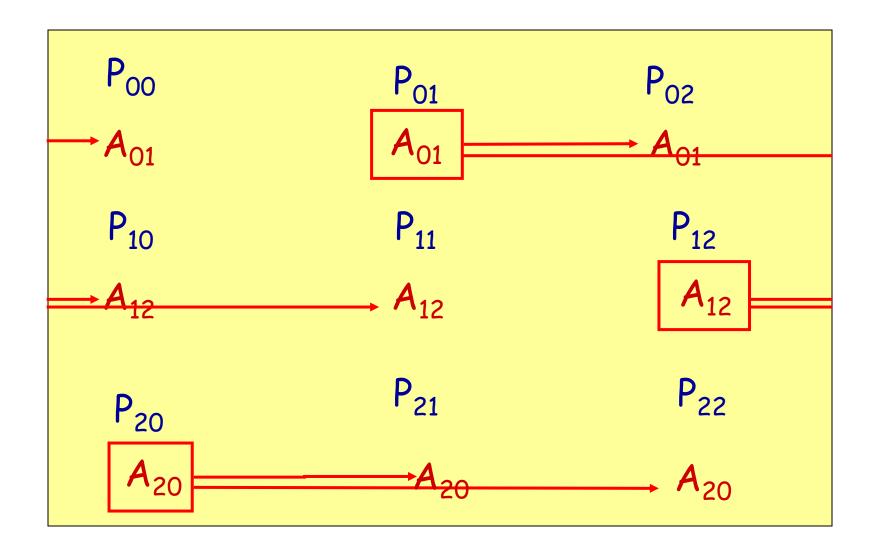
P₀₁, P₁₂ e P₂₀

inviano il proprio blocco di A,

ovvero A_{01} , A_{12} e A_{20}

Rispettivamente ai processori della riga 0,1,2 (ovvero ai processori che si trovano sulla loro stessa riga!)

II passo: invio dei blocchi di A



Osservazione: per effettuare i prodotti ...

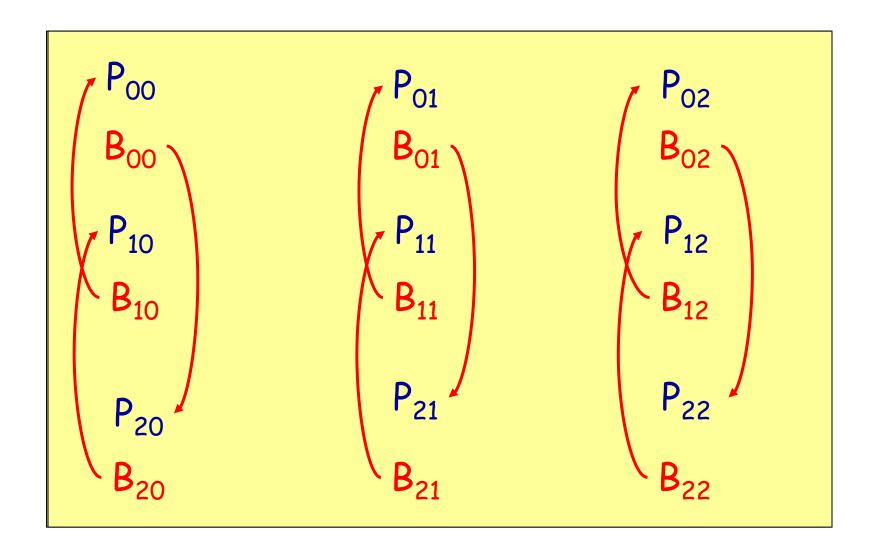
Ciascun processore ha bisogno anche del corrispondente blocco di B! (E precisamente del blocco di B

del processore nella sua stessa colonna ma nella riga successiva!) Parallel and Distributed Computing - a.a. 2023/2024

IDEA!

Per ogni colonna della griglia di processori, ciascun processore P_{ij} Invia il proprio blocco B_{ij} , al processore situato nella stessa colonna e sulla riga precedente!

Comunicazione dei blocchi di B



III passo

Nel calcolo dei C_{ij} ...

- $\cdot A_{02}$ è presente nel calcolo della I riga di C (C_{00}, C_{01}, C_{02})
- $\cdot A_{10}$ è presente nel calcolo della II riga di C (C_{10}, C_{11}, C_{12})
- G. Lace Ati, è presente nel calcolo della III riga di C (C20, C21, C22)

IDEA!

I processori:

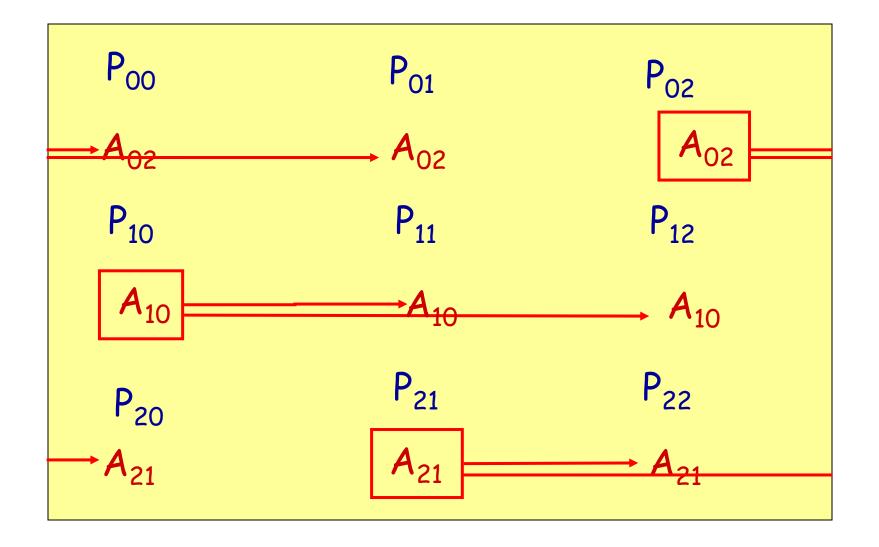
 P_{02} , P_{10} e P_{21}

inviano il proprio blocco di A,

ovvero A_{02} , A_{10} e A_{21}

Rispettivamente ai processori della riga 0,1,2 (ovvero ai processori che si trovano sulla loro stessa riga!)

II passo: invio dei blocchi di A

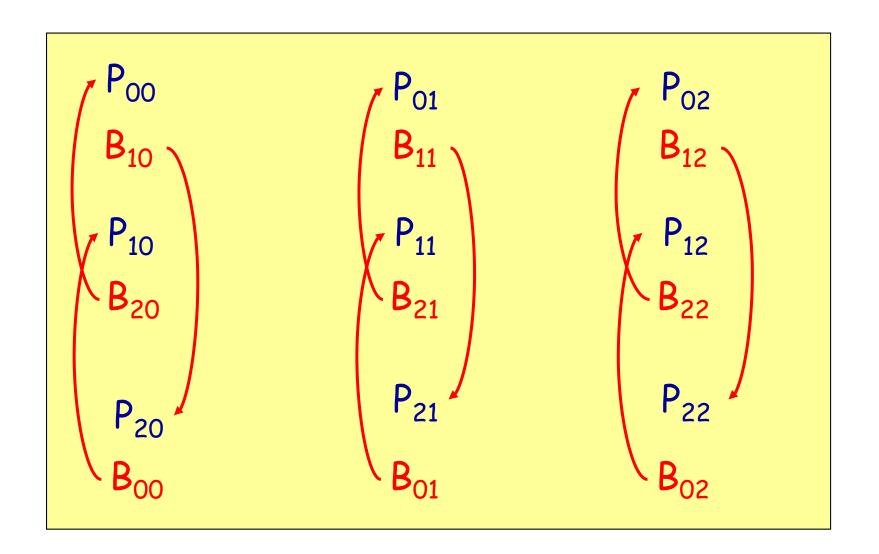


Osservazione: per effettuare i prodotti ...

Ciascun processore ha bisogno anche del corrispondente blocco di B! (E precisamente del blocco di B

del processore nella sua stessa colonna ma nella riga successiva!)

Comunicazione dei blocchi di B



$$C_{00} = A_{00} B_{00} + A_{01} B_{10} + A_{02} B_{20}$$

$$C_{01} = A_{00} B_{01} + A_{01} B_{11} + A_{02} B_{21}$$

$$C_{00} = C_{01} = C_{02} = C_{02} = C_{00} + A_{01} B_{10} + A_{02} B_{20}$$
 $A_{00} B_{01} + A_{01} B_{11} + A_{02} B_{21}$
 $A_{00} B_{02} + A_{01} B_{12} + A_{02} B_{22}$

$$C_{10}$$
 = A_{10} B_{00} + A_{11} B_{10} + A_{12} B_{20}

$$C_{11}$$
 = $A_{10} B_{01} + A_{11} B_{11} + A_{12} B_{21}$

$$C_{12}$$
 = $A_{10} B_{02} + A_{11} B_{12} + A_{12} B_{22}$

$$C_{20} =$$
 $A_{20} B_{00} + A_{21} B_{10} + A_{22} B_{20}$

$$C_{21} = A_{20} B_{01} + A_{21} B_{11} + A_{22} B_{21}$$

$$C_{22}$$
 = $A_{20} B_{02} + A_{21} B_{12} + A_{22} B_{22}$

Dopo p passi ogni processore Pii ha calcolato il corrispondente blocco C_{ii}

IV Strategia: in generale

Broadcast Multiply Rolling La strategia è costituita da p passi.

Si parte dalla diagonale principale della griglia di processori,

ad ogni passo k, si considera

la k-ma diagonale situata al di sopra di quella principale.

I processori situati lungo la diagonale effettuano una comunicazione collettiva

del blocco di A in loro possesso a tutti i processori della medesima riga.

IV Strategia: in generale

Broadcast Multiply Rolling

Inoltre, ad ogni passo
ciascun processore effettua
una comunicazione del proprio blocco di B
al processore situato
nella stessa colonna
e nella riga precedente!

Fine