ITF23019, Spring 2021

Østfold University College

Dense Matrix Multiplication

Lab 6

Outlines

- Vector-vector multiplication
- Matrix-vector multiplication
- Matrix – matrix multiplication

Vector – Vector Multiplication

Østfold University College

Vector-Vector Multiplication

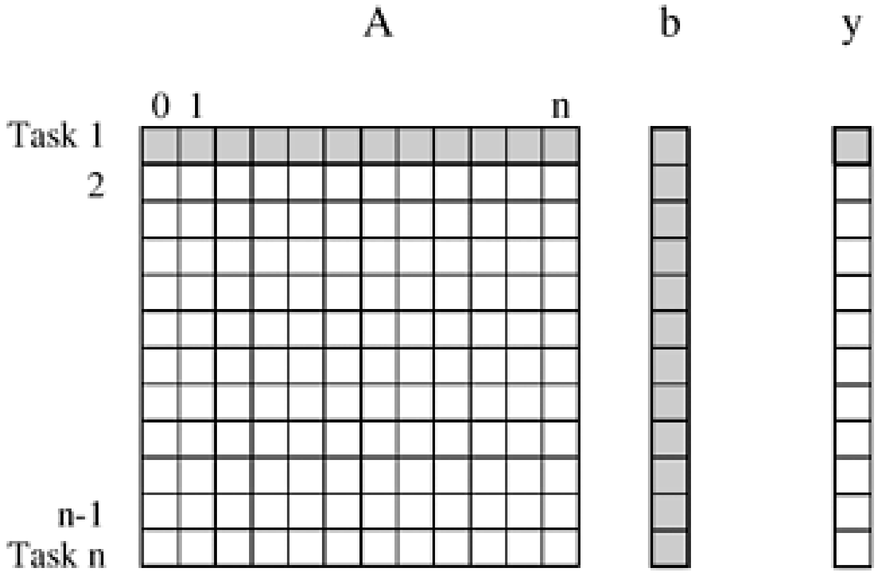
- Vector u= (u_1, u_2, u_3)
- Vector v= (u_1, u_2, u_3)
- Then, the product of vector u and v is

$$(u_1\ u_2\ u_3)\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = u_1v_1 + u_2v_2 + u_3v_3.$$

Matrix – Vector Multiplication

Matrix-Vector Multiplication

- Example



Decomposition of dense matrix-vector multiplication into n tasks, where n is the number of rows in the matrix

Serial Algorithm

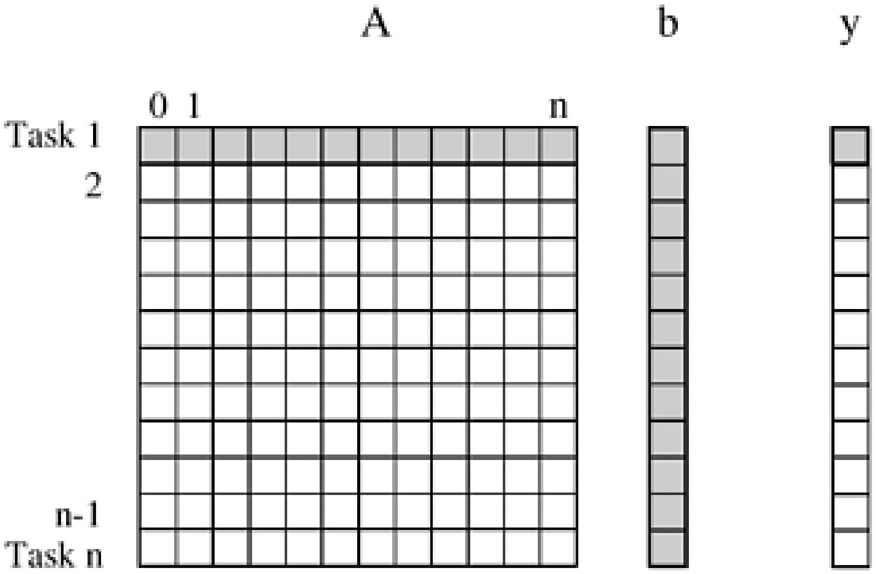
Input: Matrix A: $n \times n$
Vector b: $1 \times n$

Output: $y = A \cdot b$

Matrix y: $n \times 1$

$$y[i] = \sum_{j=1}^{n} A[i, j].b[j]$$

All tasks are independent



Decomposition of dense matrix-vector multiplication into n tasks, where n is the number of rows in the matrix

Serial Algorithm

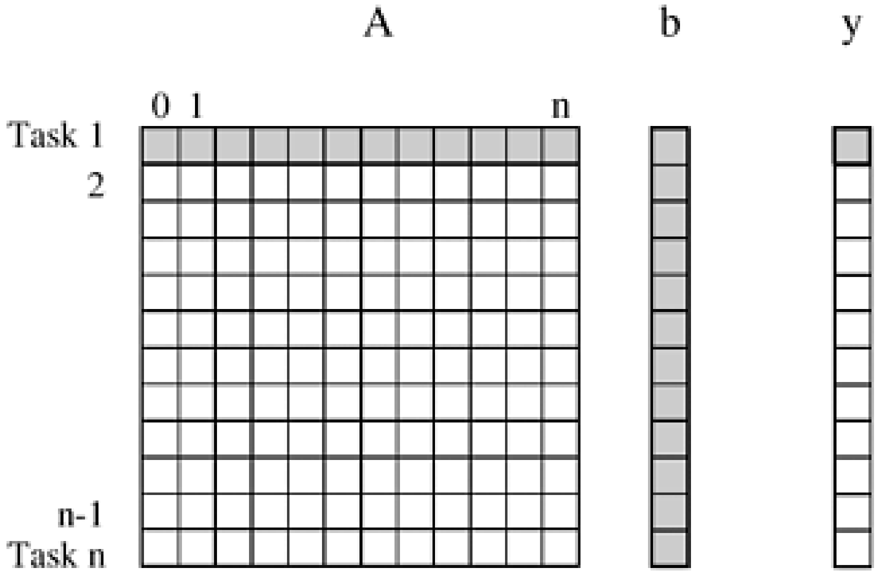
```
procedure MAT_VECT ( A, x, y)
begin
    for i := 0 to n - 1 do
    begin
        y[i]:=0;
        for j := 0 to n - 1 do
            y[i] := y[i] + A[i, j] x x[j];
    endfor;
end MAT_VECT
```

Run time/ Algorithm Complexity

$\Theta(N^2)$

Parallelism

- There are different ways of decompositions



Decomposition of dense matrix-vector multiplication into n tasks, where n is the number of rows in the matrix

Matrix – Matrix Multiplication

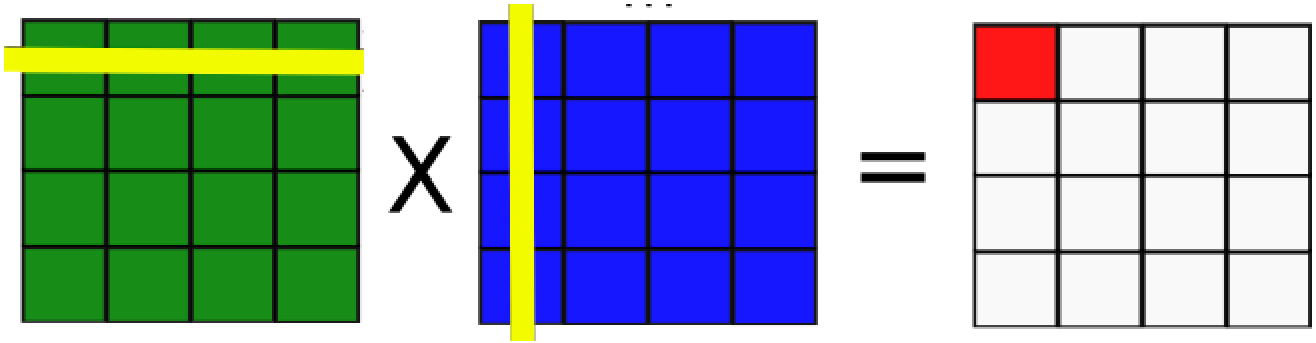
General case

- Matrix A= [a_{ij}]: size *m x n*
- Matrix B= [b_{ij}]: size *n x k*
- Product of C= [c_{ij}]: size *m x k* where c_{ij} is the product of the *i-th* row of A and the *j-th* column of B, C=AB

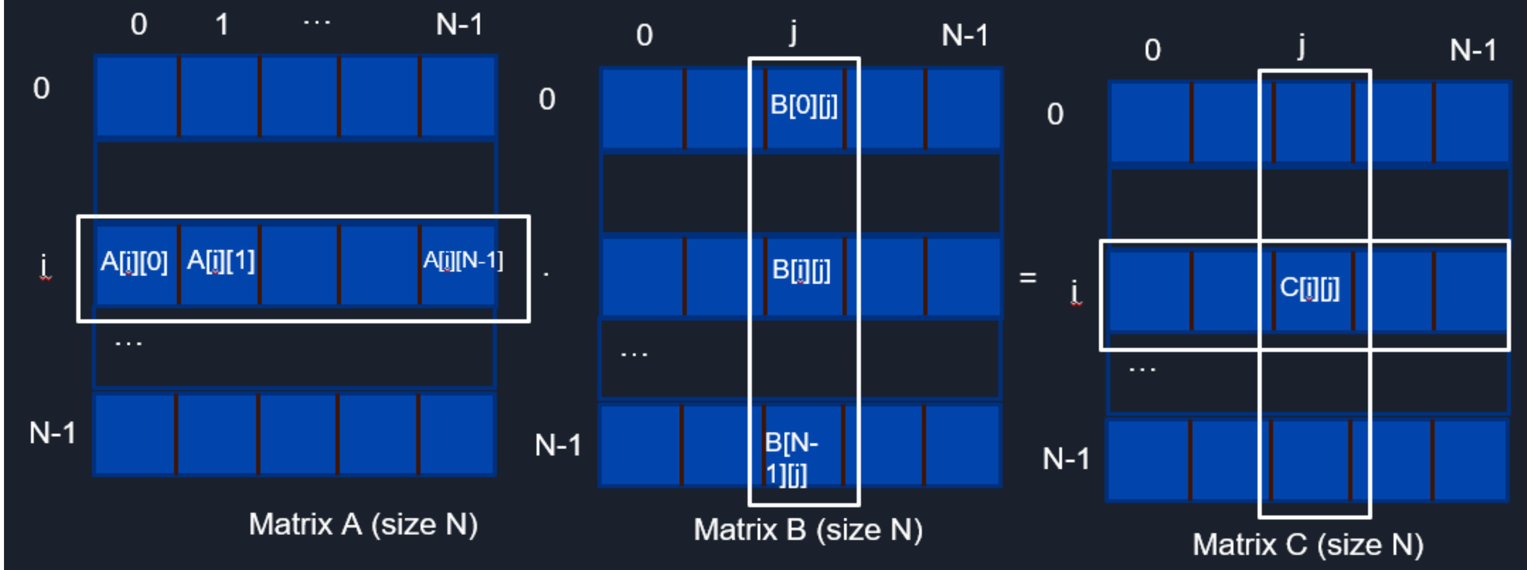
$$
\underset{A}{\begin{array}{cc} i & \begin{bmatrix} \square & \square & \square \end{bmatrix} \end{array}}
\underset{B}{\begin{array}{c} j \\ \begin{bmatrix} \square \\ \square \\ \square \end{bmatrix} \end{array}}
=
\underset{C=AB}{\begin{bmatrix} c_{ij} \end{bmatrix}}
$$

Special case

- Matrices are at the same size
- M = N = K
- Algorithm complexity: $\Theta(N^3)$



Special case



C = AB

C[i][j] = A[i][0]* B[0][j] + A[i][1]* B[1][j] +…+ A[i][N-1]* B[N-1][j]

$=\sum_{k=0}^{N-1}$ A[i][k]$*$ B[k][j]

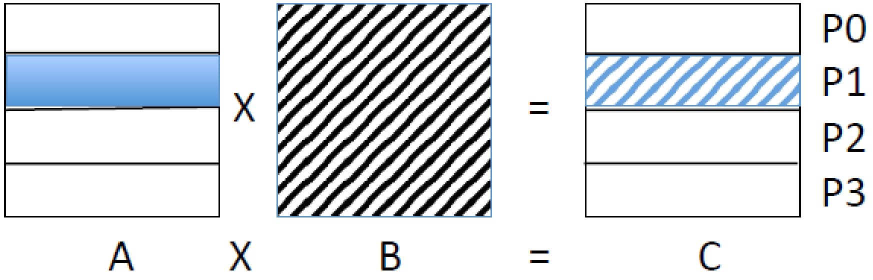
Example

$$
\begin{bmatrix} A(11) & A(12) & A(13) \\ A(21) & A(22) & A(23) \\ A(31) & A(32) & A(33) \end{bmatrix} * \begin{bmatrix} B(11) & B(12) & B(13) \\ B(21) & B(22) & B(23) \\ B(31) & B(32) & B(33) \end{bmatrix} = \begin{bmatrix} C(11) & C(12) & C(13) \\ C(21) & C(22) & C(23) \\ C(31) & C(32) & C(33) \end{bmatrix}
$$

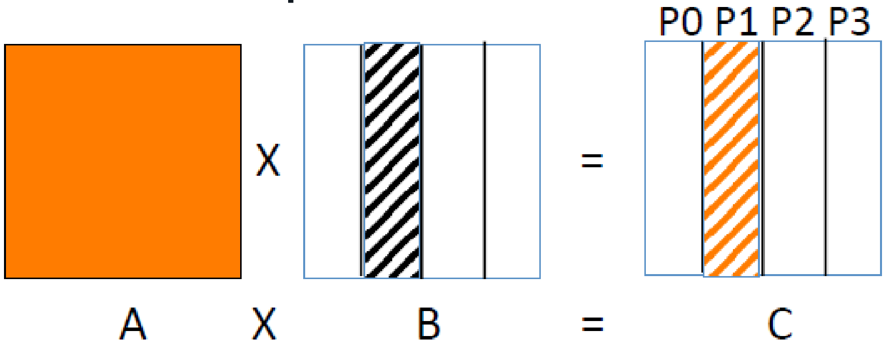
```
C(11) = A(11)*B(11) + A(12)*B(21) + A(13)*B(31)
C(21) = A(21)*B(11) + A(22)*B(21) + A(23)*B(31)
C(31) = A(31)*B(11) + A(32)*B(21) + A(33)*B(31)

C(12) = A(11)*B(12) + A(12)*B(22) + A(13)*B(32)
C(22) = A(21)*B(12) + A(22)*B(22) + A(23)*B(32)
C(32) = A(31)*B(12) + A(32)*B(22) + A(33)*B(32)

C(13) = A(11)*B(13) + A(12)*B(23) + A(13)*B(33)
C(23) = A(21)*B(13) + A(22)*B(23) + A(23)*B(33)
C(33) = A(31)*B(13) + A(32)*B(23) + A(33)*B(33)
```

Parallelism - Data Decomposition and Data Sharing

- 1-D Row-based decomposition



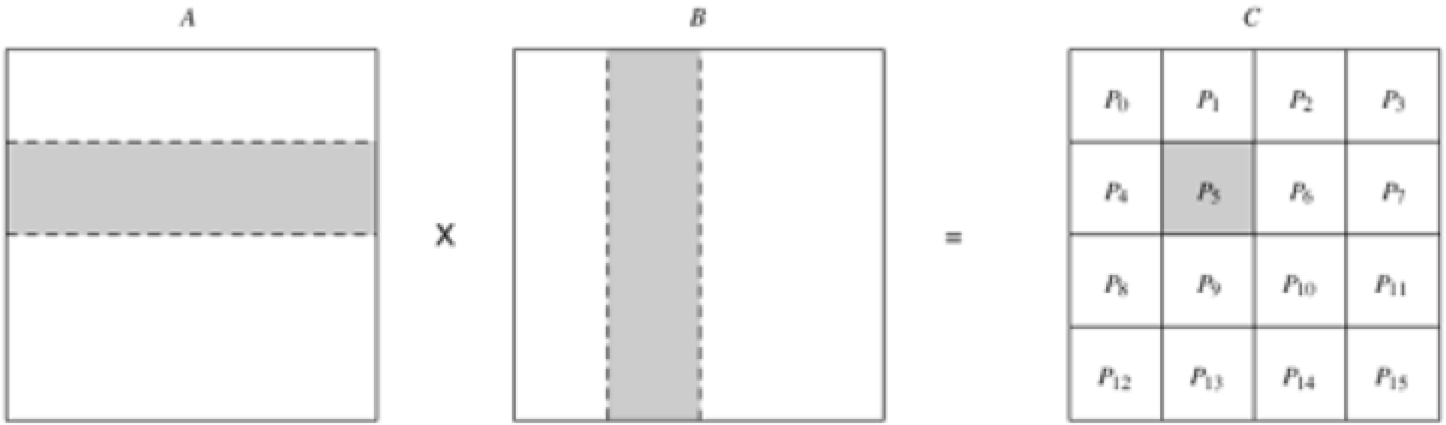
- 1-D Column-based decomposition



Østfold University College

Parallelism - Data Decomposition and Data Sharing

- 2-D Row/column-based decomposition (block distribution)



Example

$$\begin{bmatrix} A(11) & A(12) & A(13) \\ A(21) & A(22) & A(23) \\ A(31) & A(32) & A(33) \end{bmatrix} * \begin{bmatrix} B(11) & B(12) & B(13) \\ B(21) & B(22) & B(23) \\ B(31) & B(32) & B(33) \end{bmatrix} = \begin{bmatrix} C(11) & C(12) & C(13) \\ C(21) & C(22) & C(23) \\ C(31) & C(32) & C(33) \end{bmatrix}$$

```
C(11) = A(11)*B(11) + A(12)*B(21) + A(13)*B(31)
C(21) = A(21)*B(11) + A(22)*B(21) + A(23)*B(31)
C(31) = A(31)*B(11) + A(32)*B(21) + A(33)*B(31)
```

Thread 1

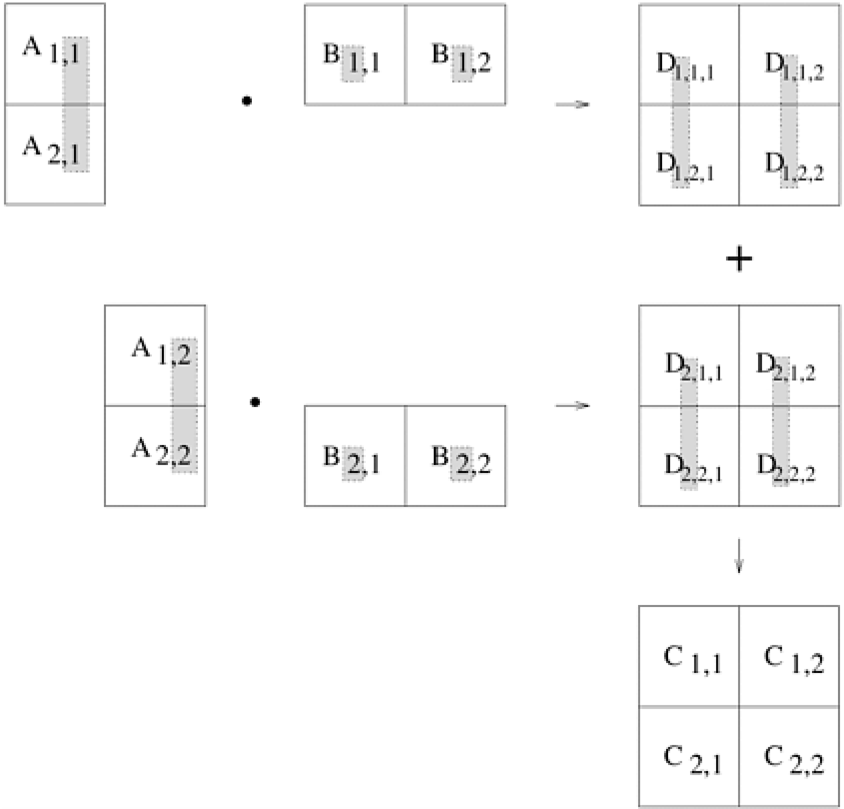
Thread 2

```
C(12) = A(11)*B(12) + A(12)*B(22) + A(13)*B(32)
C(22) = A(21)*B(12) + A(22)*B(22) + A(23)*B(32)
C(32) = A(31)*B(12) + A(32)*B(22) + A(33)*B(32)
```

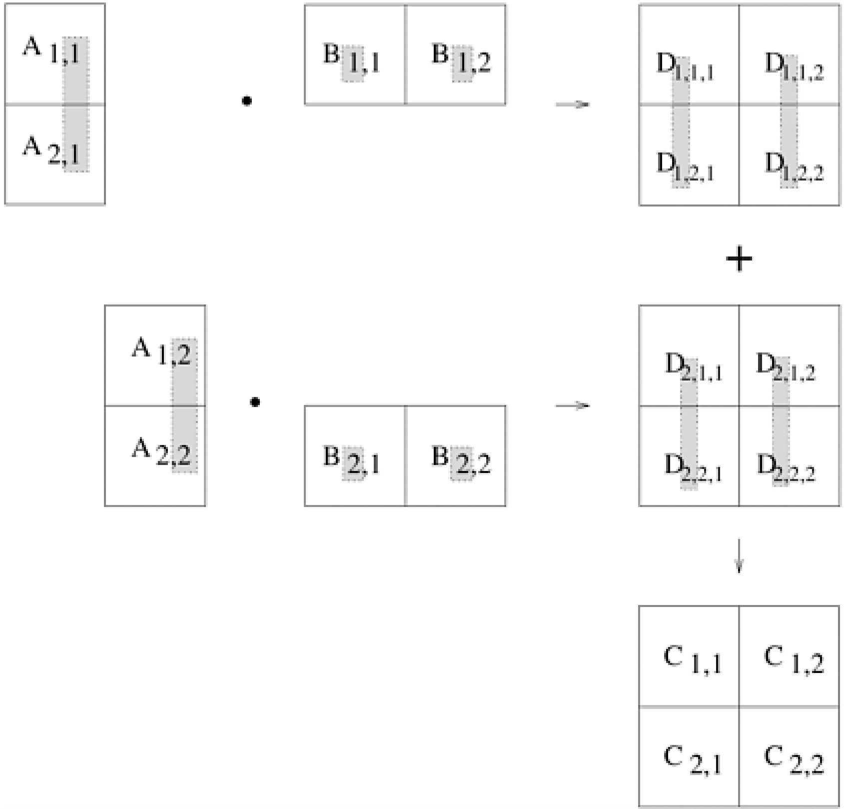
```
C(13) = A(11)*B(13) + A(12)*B(23) + A(13)*B(33)
C(23) = A(21)*B(13) + A(22)*B(23) + A(23)*B(33)
C(33) = A(31)*B(13) + A(32)*B(23) + A(33)*B(33)
```

Thread 3

Example with 3D



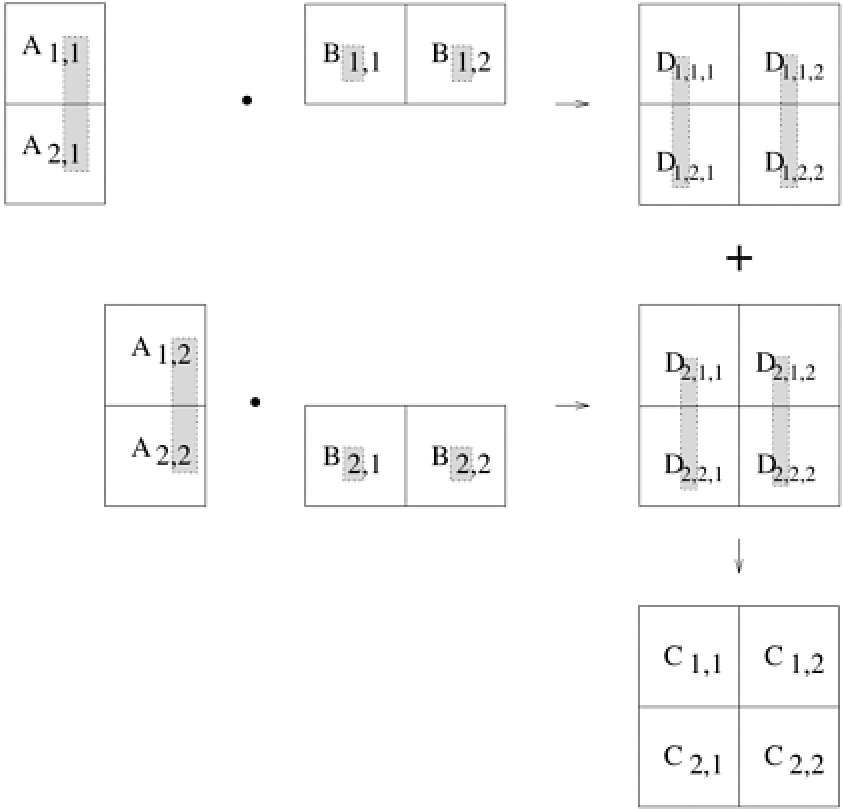
Example with 3D



Stage I

$$\begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix} \cdot \begin{pmatrix} B_{1,1} & B_{1,2} \\ B_{2,1} & B_{2,2} \end{pmatrix} \rightarrow \left(\begin{matrix} \begin{pmatrix} D_{1,1,1} & D_{1,1,2} \\ D_{1,2,2} & D_{1,2,2} \end{pmatrix} \\ \begin{pmatrix} D_{2,1,1} & D_{2,1,2} \\ D_{2,2,2} & D_{2,2,2} \end{pmatrix} \end{matrix} \right)$$

Example with 3D



Stage I

$$\begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix} \cdot \begin{pmatrix} B_{1,1} & B_{1,2} \\ B_{2,1} & B_{2,2} \end{pmatrix} \rightarrow \left(\begin{Bmatrix} \begin{pmatrix} D_{1,1,1} & D_{1,1,2} \\ D_{1,2,2} & D_{1,2,2} \end{pmatrix} \\ \begin{pmatrix} D_{2,1,1} & D_{2,1,2} \\ D_{2,2,2} & D_{2,2,2} \end{pmatrix} \end{Bmatrix} \right)$$

Stage II

$$\begin{pmatrix} D_{1,1,1} & D_{1,1,2} \\ D_{1,2,2} & D_{1,2,2} \end{pmatrix} + \begin{pmatrix} D_{2,1,1} & D_{2,1,2} \\ D_{2,2,2} & D_{2,2,2} \end{pmatrix} \rightarrow \begin{pmatrix} C_{1,1} & C_{1,2} \\ C_{2,1} & C_{2,2} \end{pmatrix}$$

Example with 3D

A decomposition induced by a partitioning of D

$$
\begin{array}{ll}
\text{Task 01:} & D_{1,1,1} = A_{1,1}B_{1,1} \\
\text{Task 02:} & D_{2,1,1} = A_{1,2}B_{2,1} \\
\text{Task 03:} & D_{1,1,2} = A_{1,1}B_{1,2} \\
\text{Task 04:} & D_{2,1,2} = A_{1,2}B_{2,2} \\
\text{Task 05:} & D_{1,2,1} = A_{2,1}B_{1,1} \\
\text{Task 06:} & D_{2,2,1} = A_{2,2}B_{2,1} \\
\text{Task 07:} & D_{1,2,2} = A_{2,1}B_{1,2} \\
\text{Task 08:} & D_{2,2,2} = A_{2,2}B_{2,2} \\
\text{Task 09:} & C_{1,1} = D_{1,1,1} + D_{2,1,1} \\
\text{Task 10:} & C_{1,2} = D_{1,1,2} + D_{2,1,2} \\
\text{Task 11:} & C_{2,1} = D_{1,2,1} + D_{2,2,1} \\
\text{Task 12:} & C_{2,2} = D_{1,2,2} + D_{2,2,2}
\end{array}
$$

Stage I

$$
\begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix} \cdot \begin{pmatrix} B_{1,1} & B_{1,2} \\ B_{2,1} & B_{2,2} \end{pmatrix} \rightarrow \left(\begin{array}{c} \begin{pmatrix} D_{1,1,1} & D_{1,1,2} \\ D_{1,2,2} & D_{1,2,2} \end{pmatrix} \\ \begin{pmatrix} D_{2,1,1} & D_{2,1,2} \\ D_{2,2,2} & D_{2,2,2} \end{pmatrix} \end{array} \right)
$$

Stage II

$$
\begin{pmatrix} D_{1,1,1} & D_{1,1,2} \\ D_{1,2,2} & D_{1,2,2} \end{pmatrix} + \begin{pmatrix} D_{2,1,1} & D_{2,1,2} \\ D_{2,2,2} & D_{2,2,2} \end{pmatrix} \rightarrow \begin{pmatrix} C_{1,1} & C_{1,2} \\ C_{2,1} & C_{2,2} \end{pmatrix}
$$

Example with 3D

A decomposition induced by a partitioning of D

Task 01: $D_{1,1,1} = A_{1,1}B_{1,1}$

Task 02: $D_{2,1,1} = A_{1,2}B_{2,1}$

Task 03: $D_{1,1,2} = A_{1,1}B_{1,2}$

Task 04: $D_{2,1,2} = A_{1,2}B_{2,2}$

Task 05: $D_{1,2,1} = A_{2,1}B_{1,1}$

Task 06: $D_{2,2,1} = A_{2,2}B_{2,1}$

Task 07: $D_{1,2,2} = A_{2,1}B_{1,2}$

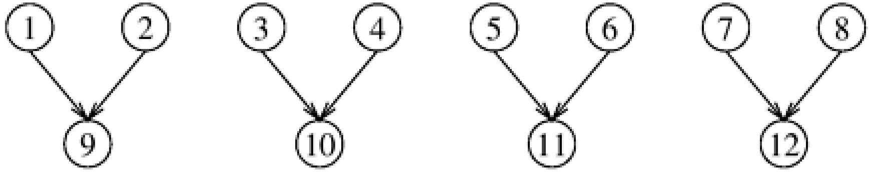
Task 08: $D_{2,2,2} = A_{2,2}B_{2,2}$

Task 09: $C_{1,1} = D_{1,1,1} + D_{2,1,1}$

Task 10: $C_{1,2} = D_{1,1,2} + D_{2,1,2}$

Task 11: $C_{2,1} = D_{1,2,1} + D_{2,2,1}$

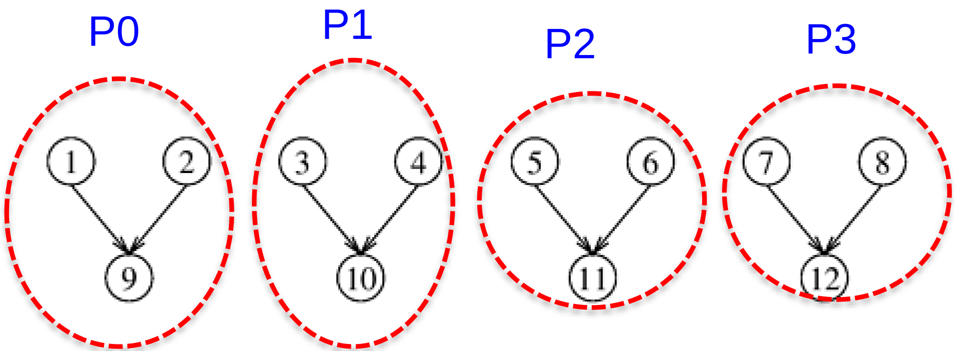
Task 12: $C_{2,2} = D_{1,2,2} + D_{2,2,2}$



Task dependence graph

Østfold University College

Example with 3D



Task dependence graph

References

[1] Introduction to Parallel Computing, Ananth Grama, George Karypis, Vipin Kumar, Anshul Gupta, Addison Wesley, 2003, Chapter 9.