MPI implementation of a PCG solver for nonconforming FEM problems: overlapping of communications and computations

Gergana Bencheva, Svetozar Margenov

Institute for Parallel Processing, Bulgarian Academy of Sciences

gery@parallel.bas.bg, margenov@parallel.bas.bg

Jiří Starý

Institute of Geonics, Academy of Sciences of Czech Republic

stary@ugn.cas.cz

Preliminaries

Formulation of the problem

Consider the second order elliptic boundary value problem:

$$(1)\quad \left| \begin{array}{rcll} -\nabla \cdot (a(y)\nabla u(y)) & = & f(y), & y = (y_1, y_2) \in \Omega \subset \mathbb{R}^2, \\ u & = & \mu(y), & y \in \Gamma_D \ (meas(\Gamma_D) \neq 0), \\ (a(y)\nabla u(y)) \cdot n & = & g(y), & y \in \Gamma_N. \end{array} \right.$$

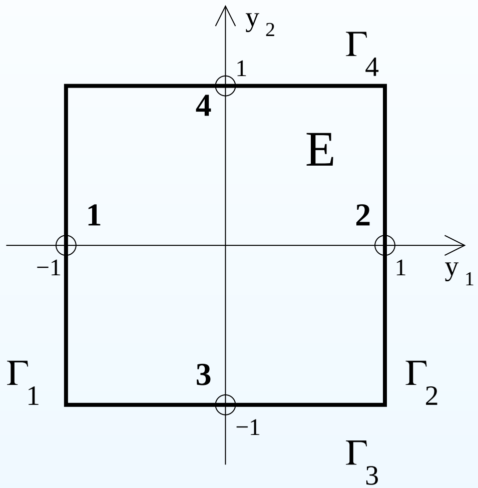
$$f(y),\ \mu(y),\ g(y) \in L^2(\Omega),\ \Gamma = \partial\Omega = \Gamma_D \cup \Gamma_N,$$
$$a(y) = \{a_{ij}(y)\}_{i,j=1}^2$$

$\Downarrow$ Discretization (FDM, FEM)

$$A\mathbf{x} = \mathbf{f}$$

Goal: *Scalable Parallel Preconditioner*

Non-conforming quadrilateral finite elements



Reference element E.

Basis functions $\widehat{\phi}_i \in S_p$,

$$S_p = span\{1, y_1, y_2, y_1^2 - y_2^2\}$$

MP: $\widehat{\phi}_i(j) = \delta_{i,j},\ i, j = 1, \ldots, 4$

MV: $\dfrac{1}{|\Gamma_j|} \displaystyle\int_{\Gamma_j} \widehat{\phi}_i\, dy = \delta_{i,j},\ i, j = 1, \ldots, 4$

Basis MP

$$\widehat{\phi}_1(y_1, y_2) = \tfrac{1}{4}(1 - 2y_1 + (y_1^2 - y_2^2))$$
$$\widehat{\phi}_2(y_1, y_2) = \tfrac{1}{4}(1 + 2y_1 + (y_1^2 - y_2^2))$$
$$\widehat{\phi}_3(y_1, y_2) = \tfrac{1}{4}(1 - 2y_2 - (y_1^2 - y_2^2))$$
$$\widehat{\phi}_4(y_1, y_2) = \tfrac{1}{4}(1 + 2y_2 - (y_1^2 - y_2^2))$$

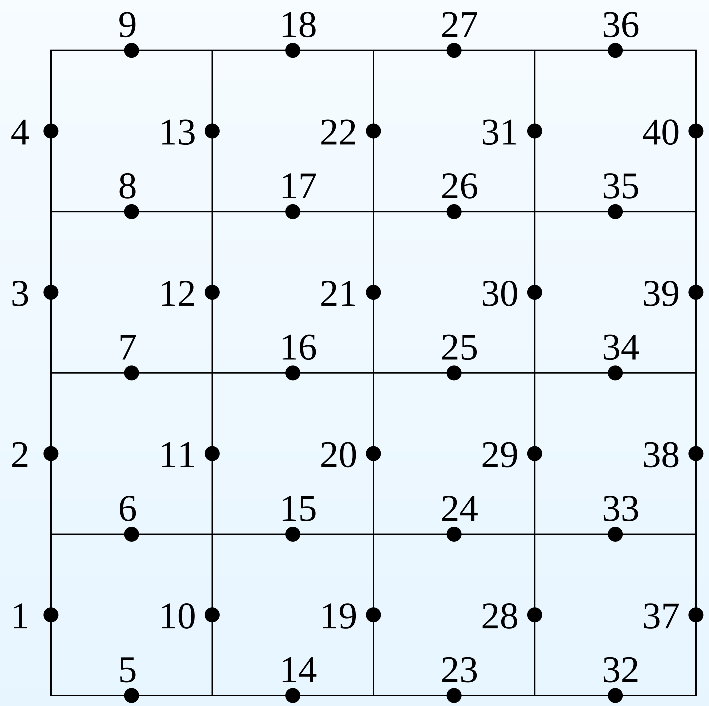
Basis MV

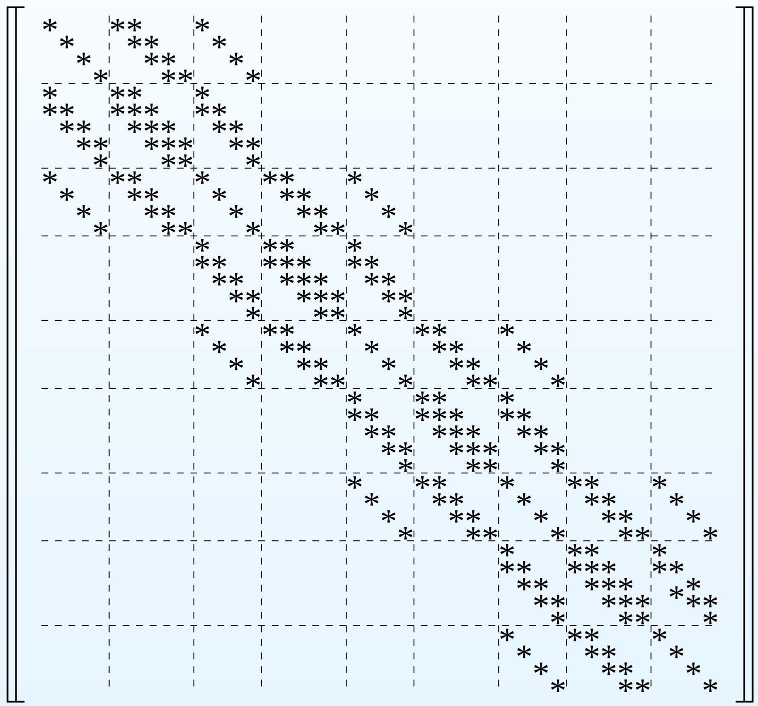
$$\widehat{\phi}_1(y_1, y_2) = \tfrac{1}{8}(2 - 4y_1 + 3(y_1^2 - y_2^2))$$
$$\widehat{\phi}_2(y_1, y_2) = \tfrac{1}{8}(2 + 4y_1 + 3(y_1^2 - y_2^2))$$
$$\widehat{\phi}_3(y_1, y_2) = \tfrac{1}{8}(2 - 4y_2 - 3(y_1^2 - y_2^2))$$
$$\widehat{\phi}_4(y_1, y_2) = \tfrac{1}{8}(2 + 4y_2 - 3(y_1^2 - y_2^2))$$

Structure of the stiffness matrix

Non-conforming quadrilateral elements

$n_1 \times n_2$ mesh, $N = n_1(2n_2 + 1) + n_2$





a) nodes' numbering; b) stiffness matrix $A_{N \times N}$

Why nonconforming FEM?

- better approximation for some ill conditioned problems
  - Stokes problem (R. Rannacher and S. Turek (1992))
  - Elasticity problem in the case of almost incompressible materials (P. Hansbo and M. Larson (2001))
- regular sparsity structure of the stiffness matrix for non-regular mesh
- specifi c opportunities for parallel implementation

Background Solution Method

Preconditioned Conjugate Gradient (PCG) with
Modifi ed Incomplete Cholesky (MIC(0)) Preconditioner

$$A = D - L - L^t\,, \quad X = diag(x_1, \ldots, x_N)$$

$$L \geq 0\,, \quad A\underline{e} \geq 0\,, \quad A\underline{e} + L^t\underline{e} > 0\,, \quad \underline{e} = (1, \cdots, 1)^t \in \mathcal{R}^{\mathcal{N}}\,,$$

$$x_i = a_{ii} - \sum_{k=1}^{i-1} \frac{a_{ik}}{x_k} \sum_{j=k+1}^{N} a_{kj}\,, \quad x_i > 0\,.$$

$$\Downarrow$$

$$\mathcal{C}_{MIC(0)}(A) = (X - L)X^{-1}(X - L)^t$$

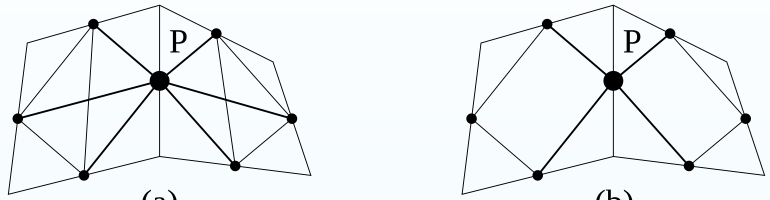
is stable MIC(0) factorization of A.

Preconditioning Strategy

Preconditioning Strategy

1) Local modifi cation known as „diagonal compensation“:

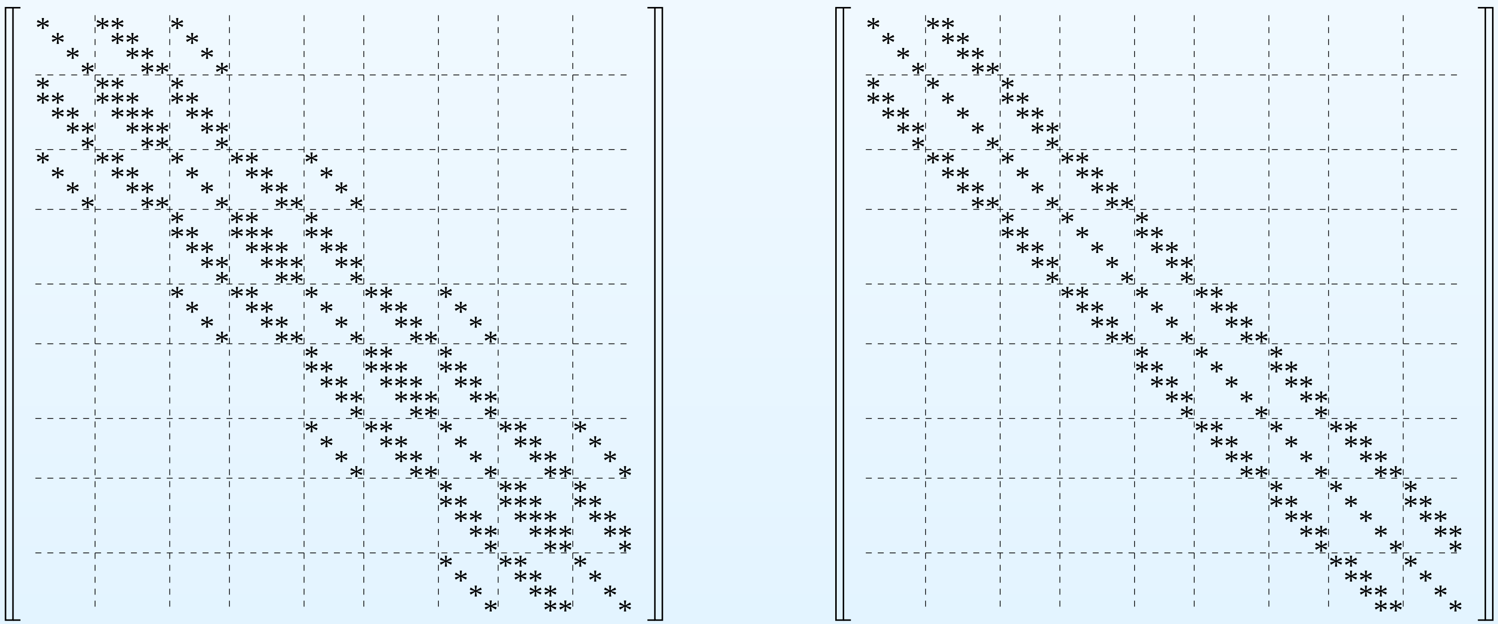
$$A \to B;$$



(a) (b)

2) $\mathrm{MIC}(0)$ factorization: preconditioner $\mathcal{C} = \mathcal{C}_{\mathrm{MIC}(0)}(B)$ for A

Structure of the matrix A and the introduced matrix B



(a) (b)

Convergence rate

Theorem

(i) the sparse approximation B of the stiffness matrix A satisfi es the conditions for a stable $\mathrm{MIC}(0)$ factorization;

(ii) the matrices B and A are spectrally equivalent where the next relative condition number estimate holds uniformly with respect to any possible coeffi cients jumps:

$$\kappa((B^{MP})^{-1}A^{MP}) \leq 2 \quad \text{for} \quad \varepsilon \in [\frac{1}{2}, 1]\,,$$

$$\kappa((B^{MV})^{-1}A^{MV}) \leq 3 \quad \text{for} \quad \varepsilon \in [\frac{1}{3}, 1]\,.$$

$\kappa(\mathcal{C}^{-1}A) = \mathcal{O}(N^{\frac{1}{2}})$, where $\mathcal{C} = \mathcal{C}_{\mathrm{MIC}(0)}(B)$

$$\mathcal{N}_{it}^{PCG/\mathrm{MIC}(0)}(A^{-1}\mathbf{b}) \approx 34\, N$$

Parallel Implementation

Parallel implementation

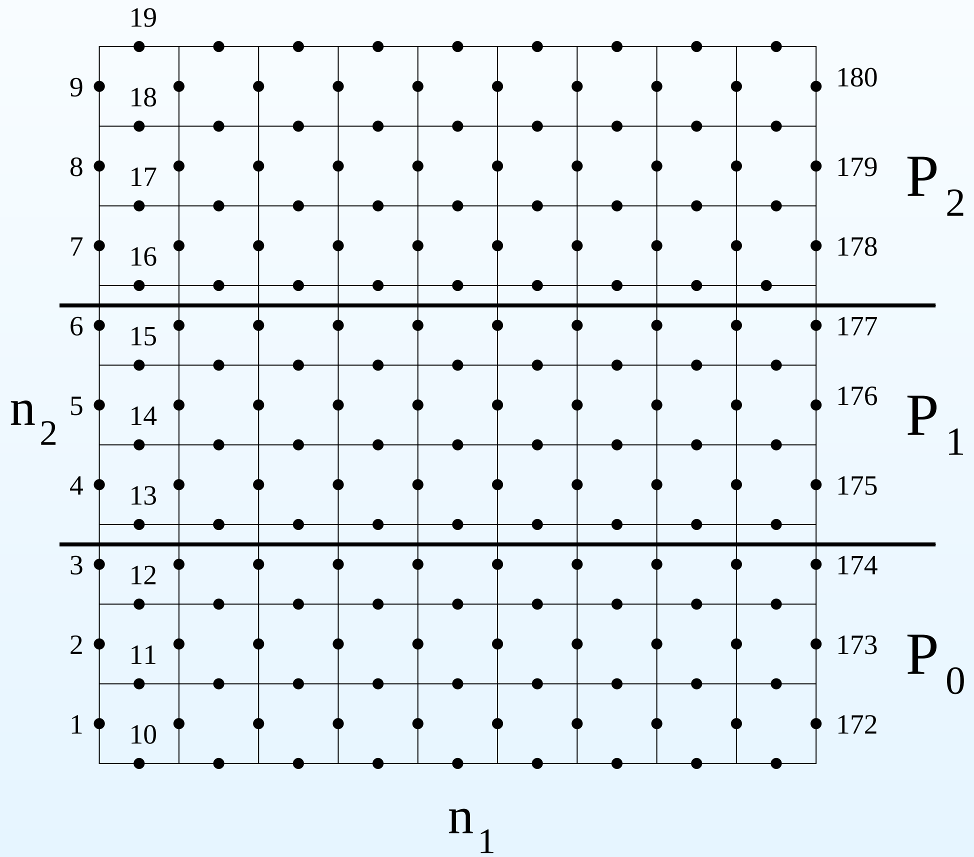
$N = n_1(2n_2+1) + n_2$, N_p – number of processors

1. Data – the domain is partitioned into N_p horizontal strips
2. Computations – equally distributed among the processors
   - 1 solution of system with $\mathcal{C}_{N\times N}$ ($\approx 11N/N_p$ a. o.)
   - 1 matrix vector multiplication with $A_{N\times N}$ ($\approx 13N/N_p$ a. o.)
   - 2 inner products ($4N/N_p$ a. o.)
   - 3 linked vector triads $\mathbf{v} := \alpha\mathbf{v} + \mathbf{u}$ ($6N/N_p$ a. o.)
3. Communications
   - inner products – global;
   - matrix-vector multiplication – local;
   - system with $\mathcal{C}$ – local;

Each block equation is handled in parallel.

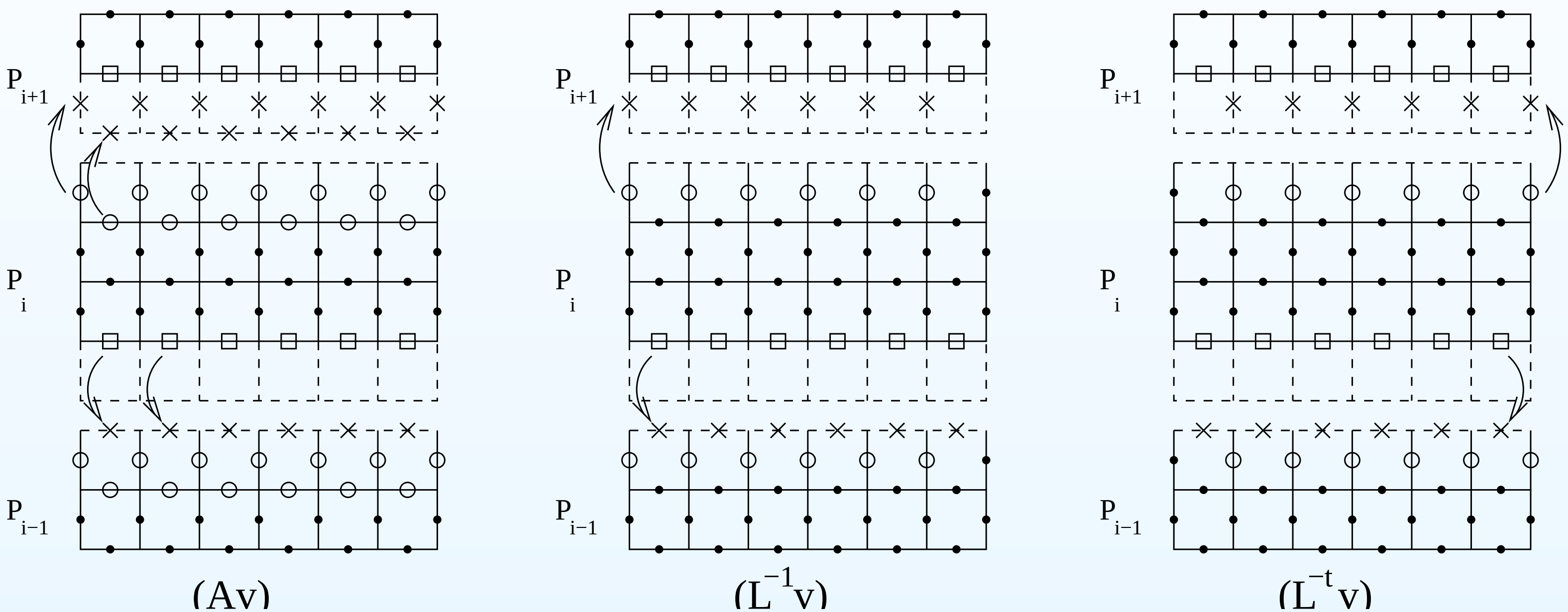
$$T_{N_p}^{it} = T_a^{it} + T_{com}^{it} \approx 34\frac{n_1(2n_2+1)+n_2}{N_p}.t_a + 8n_1.t_s + 14n_1.t_w$$

Data Partitioning

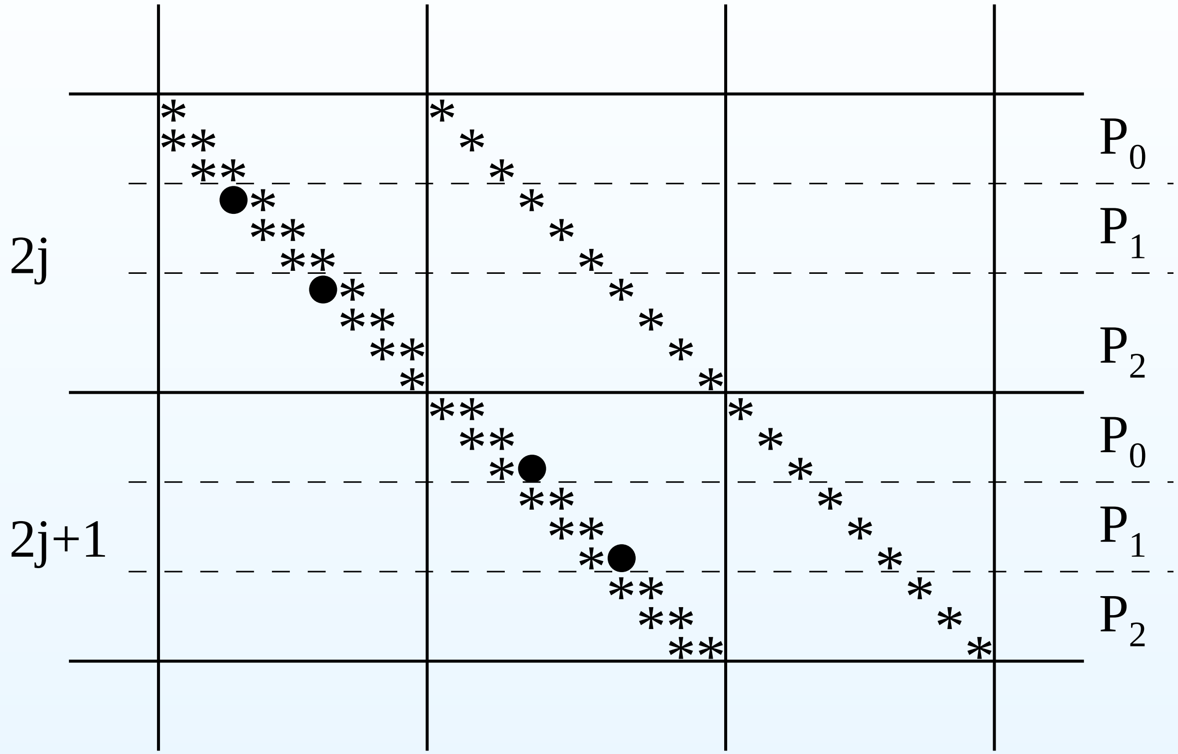


$$N = n_1(2n_2 + 1) + n_2\,,\; N_p = 3\,,\; n_1 = 9\,,\; n_2 = 9\,,$$

Communications



Overlapping of communications and computations



$$\mathcal{C}_{\mathrm{MIC}(0)}(B)\mathbf{w} \equiv (X - \widetilde{L})X^{-1}(X - \widetilde{L})^t\mathbf{w} = \mathbf{v}$$

1) find $\mathbf{y}$ from $L\mathbf{y} = \mathbf{v}$, where $L = X - \widetilde{L}$;
2) compute $\mathbf{y} := X\mathbf{y}$ (no communications are required);
3) find $\mathbf{w}$ from $L^t\mathbf{w} = \mathbf{y}$.

Numerical Tests

Parallel computing systems

Thea – cluster of 8 computers, each with 1.5 GB of RAM and a single AMD Athlon processor at 1.4GHz (Institute of Geonics, Academy of Sciences of Czech Republic, Ostrava, Czech Republic).

Simba – separate domain of a Sun Fire 15k server with 36 UltraSPARC III+ CPUs at 900 MHz and 36 GB of RAM (Department of Information Technology, Uppsala University, Sweden)

Algorithm MP

		Thea						Simba					
		noverlap			overlap			noverlap			overlap		
N_p	$\frac{n}{iter}$	cpu	S_{N_p}	E_{N_p}	cpu	S_{N_p}	E_{N_p}	cpu	S_{N_p}	E_{N_p}	cpu	S_{N_p}	E_{N_p}
1		9.16			9.21			16.36			16.36		
2	$\frac{256}{71}$	9.55	0.96	0.48	8.10	1.14	0.57	7.97	2.05	1.03	7.71	2.12	1.06
4		11.51	0.80	0.20	7.41	1.24	0.31	3.39	4.83	1.21	3.20	5.11	1.28
8		11.20	0.82	0.10	6.47	1.37	0.17	2.48	6.60	0.82	2.43	6.73	0.84
16								3.24	5.05	0.32	2.88	5.68	0.36
1		54.11			54.22			108.11			108.07		
2	$\frac{512}{104}$	41.91	1.29	0.65	33.88	1.60	0.80	54.57	1.98	0.99	53.46	2.02	1.01
4		41.35	1.31	0.33	24.97	2.17	0.54	29.06	3.72	0.93	29.13	3.71	0.93
8		36.47	1.48	0.19	20.65	2.63	0.33	15.38	7.03	0.88	14.86	7.27	0.91
16								11.10	9.77	0.61	9.84	10.98	0.69
1		286.91			287.51			646.95			647.40		
2	$\frac{1024}{148}$	212.41	1.35	0.68	192.32	1.49	0.75	325.05	1.99	1.00	323.91	2.00	1.00
4		155.52	1.84	0.46	107.49	2.67	0.67	170.77	3.79	0.95	167.87	3.86	0.96
8		125.01	2.30	0.29	71.09	4.04	0.51	88.85	7.28	0.91	86.49	7.49	0.94
16								52.37	12.35	0.77	51.95	12.46	0.78

Algorithm MV

		Thea						Simba					
		noverlap			overlap			noverlap			overlap		
N_p	$\frac{n}{iter}$	cpu	S_{N_p}	E_{N_p}	cpu	S_{N_p}	E_{N_p}	cpu	S_{N_p}	E_{N_p}	cpu	S_{N_p}	E_{N_p}
1		10.43			10.47			18.65			18.63		
2	256	10.89	0.96	0.48	9.23	1.13	0.57	8.97	2.08	1.04	8.78	2.12	1.06
4	81	13.10	0.80	0.20	8.42	1.24	0.31	3.73	5.00	1.25	3.73	4.99	1.25
8		12.70	0.82	0.10	7.73	1.35	0.17	2.83	6.59	0.82	2.82	2.82	0.83
16								3.68	5.07	0.32	3.35	5.56	0.35
1		61.70			61.88			123.59			123.59		
2	512	47.96	1.29	0.65	38.77	1.60	0.80	61.78	2.00	1.00	61.16	2.02	1.01
4	119	47.27	1.31	0.33	28.65	2.16	0.54	33.84	3.65	0.91	33.21	3.72	0.93
8		41.31	1.49	0.19	23.61	2.62	0.33	17.35	7.12	0.89	16.83	7.34	0.92
16								12.77	9.68	0.60	10.99	11.25	0.70
1		323.42			323.77			729.48			729.87		
2	1024	239.60	1.35	0.68	216.89	1.49	0.75	366.47	1.99	1.00	365.30	2.00	1.00
4	167	175.52	1.84	0.46	121.26	2.67	0.67	196.10	3.72	0.93	189.57	3.85	0.96
8		140.60	2.30	0.29	80.06	4.04	0.51	97.74	7.46	0.93	101.75	7.17	0.90
16								58.47	12.48	0.78	58.33	12.51	0.78

Acknowledgments

This research has been supported in part by:

- the Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under the project p2004009 „Parallel computing in Geosciences“
- the bilateral IG AS – IPP BAS interacademy exchange grant „Reliable Modelling and Large Scale Computing in Geosciences“

This scientifi c visit was possible due to the sponsorship of the Royal Swedish Academy of Engineering Sciences, IVA

Thank you for your attention!