

SSOR Preconditioned Conjugate Gradient Algorithm in Dynamic SMP Clusters with Communication on the Fly

Boguslaw Butrylo



Bialystok Technical
Univeristy,
Poland

**Marek Tudruj
Lukasz Masko**



Polish Academy of
Sciences,
Poland

**Laurent Nicolas
Christian Vollaire**



Centre de Genie
Electrique de Lyon,
France

Outline

- ◆ Introduction.
- ◆ Formulation of the FEM algorithm.
- ◆ Graph of the solver: PCG+SSOR.
- ◆ Distributed formulation.
- ◆ Message passing formulation.
- ◆ Dynamic SMP clusters.
- ◆ Conclusions.

Introduction

Accuracy, fidelity of numerical approximation

Numerical simulation of electromagnetic phenomena

Time of numerical simulation

A conjugate gradient algorithm (CG) is one of the most suitable methods in time domain analysis of electromagnetic field:

- ♦ The representative computational metrics of the algorithms are obtained by the analysis of the benchmark problem both in sequential and in multicomputer message passing environments.
- ♦ The principles of the implementation of PCG algorithm with SSOR preconditioner in dynamically configurable SMP clusters are presented.

Formulation of the FEM algorithm

- Non-harmonic periodic excitations and electromagnetic impulses are analysed in time domain. The spatial, time-variable distribution of the EM field is described by the Maxwell equations

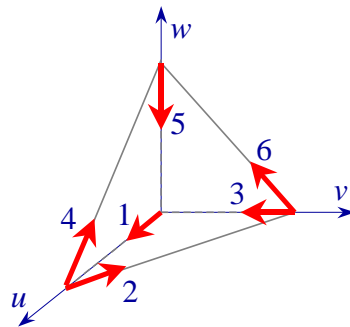
$$\text{rot } \mathbf{H} = \gamma \mathbf{E} + \varepsilon \frac{\partial \mathbf{E}}{\partial t}$$

$$\text{rot } \mathbf{E} = -\mu \frac{\partial \mathbf{H}}{\partial t}$$

$$\varepsilon \frac{\partial^2 \mathbf{e}(t)}{\partial t^2} + \sigma \frac{\partial \mathbf{e}(t)}{\partial t} + \nabla \times \frac{1}{\mu} \nabla \times \mathbf{e}(t) = 0$$

$$\mathbf{h}(t) = -\frac{1}{\mu} \int_0^t \nabla \times \mathbf{e}(t) dt$$

- The spatial discretization of the analysed model is made with edge-based finite elements



$$\mathbf{e}(t) = \sum_{i=1}^{N_E} e_i(t) \cdot \mathbf{W}_i$$

$$\mathbf{W}_k = \lambda_i \nabla \lambda_j - \lambda_j \nabla \lambda_i$$

Nedge Elements $H(\text{rot}, \Omega)$

Formulation of the FEM algorithm

- The temporal discretization

$$\left(\mathbf{T} + \frac{\Delta t}{2} \mathbf{R} \right) \cdot \mathbf{e}_{n+1} = \left(2\mathbf{T} - \Delta t^2 \mathbf{S} \right) \cdot \mathbf{e}_n + \left(\frac{\Delta t}{2} \mathbf{R} - \mathbf{T} \right) \cdot \mathbf{e}_{n-1}$$

$$t_{ij} = \int_{V_p} \varepsilon \mathbf{W}_i \mathbf{W}_j dV$$

$$r_{\sigma,ij} = \int_{V_p} \sigma \mathbf{W}_i \mathbf{W}_j dV$$

$$r_{ABC,ij} = \int_{S_{\infty,p}} \mathbf{W}_j \frac{1}{\mu c} (\mathbf{W}_i \times \bar{\mathbf{n}}) dS$$

$$s_{ij} = \int_{V_p} \frac{1}{\mu} (\nabla \times \mathbf{W}_i) (\nabla \times \mathbf{W}_j) dV$$

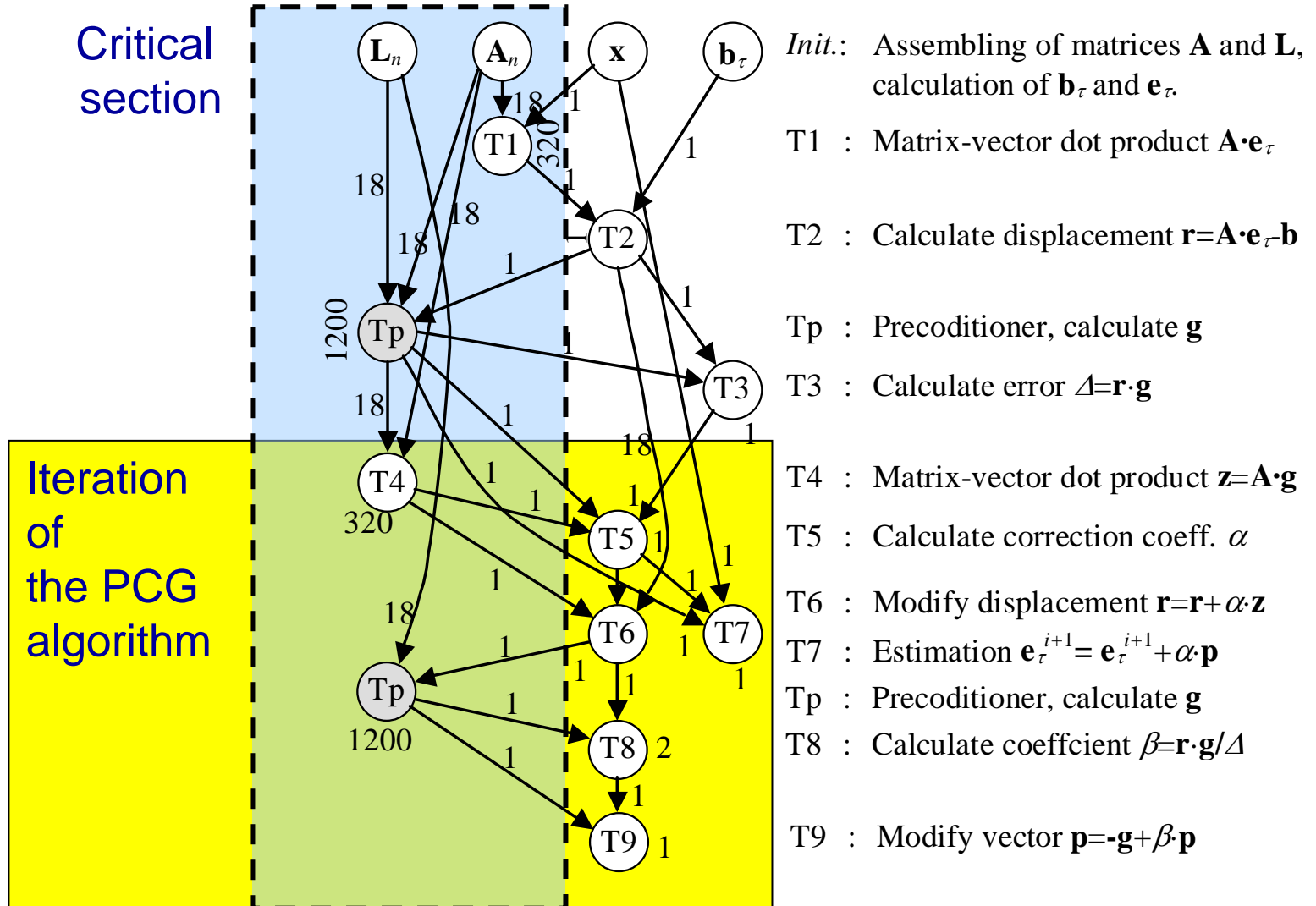
The large scale matrix equation:

$$\mathbf{A} \mathbf{e}_{n+1} = \mathbf{b}_{n+1}$$

The A matrix is:

- well-conditioned,
- symmetric,
- diagonal dominant,
- sparse,
- real,
- compressed in the CRS form.

Graph of the solver: PCG+SSOR

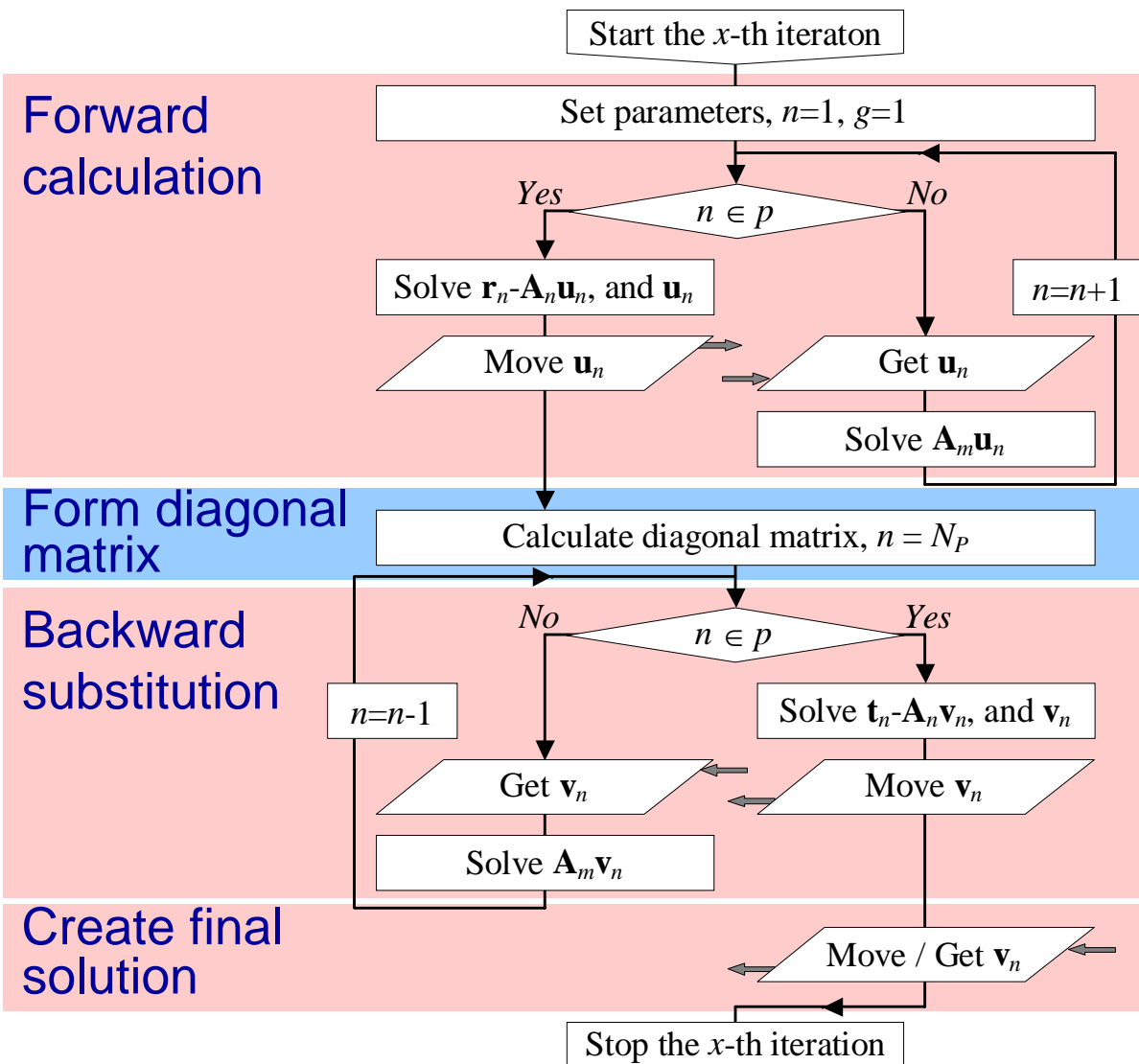


Distributed formulation: PCG+SSOR

- ◆ The presented implementation of the solver is based on the domain decomposition paradigm and the SPMD scheme.
- ◆ The A matrix is decomposed in two steps:
 - ◆ The CRS form of the matrix extorts the row-wise decomposition.
 - ◆ The locally defined sub-matrices \mathbf{A}_n are decomposed into some column-wise sub-matrices.

The sub-matrices \mathbf{A}_n (the 1-st decomposition) and then \mathbf{A}_{nm} (the 2-nd decomposition) do not overlap.

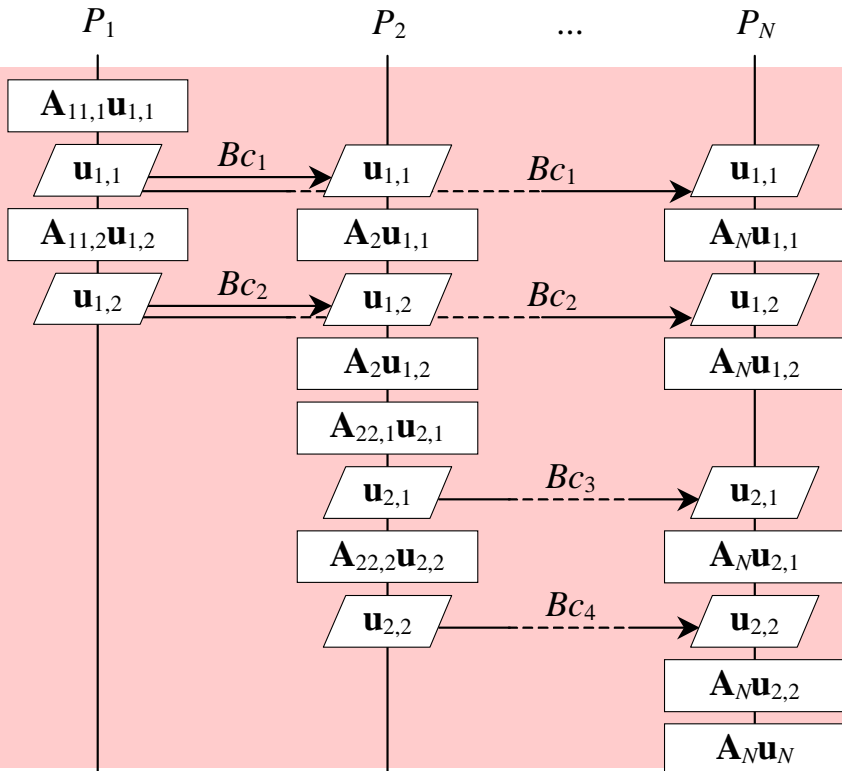
Message passing formulation



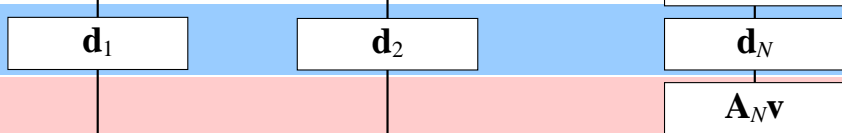
- The second level decomposition enables adjusting the computation grain.
- The granulation degree is determined by granulation coefficient G .
- Large G corresponds to a fine grain decomposition of the algorithm into smaller pieces of computations and communication. The processor performs relatively few computations between consecutive data transfers.

Message passing formulation

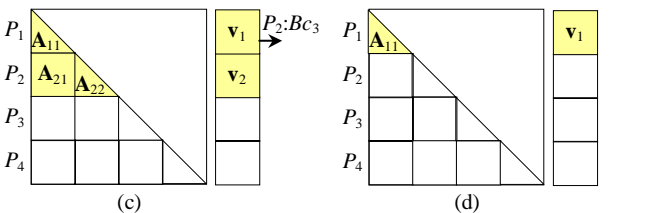
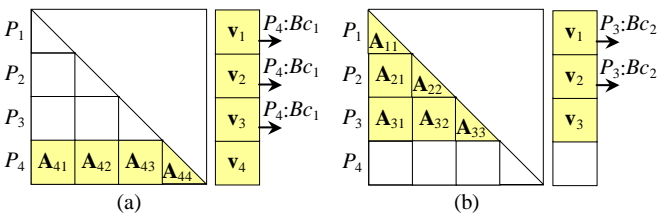
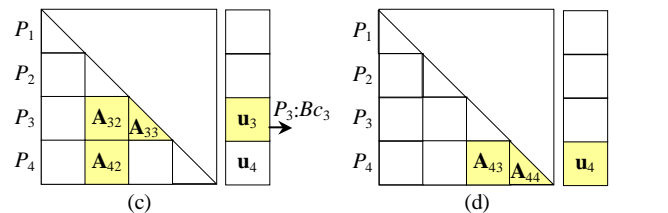
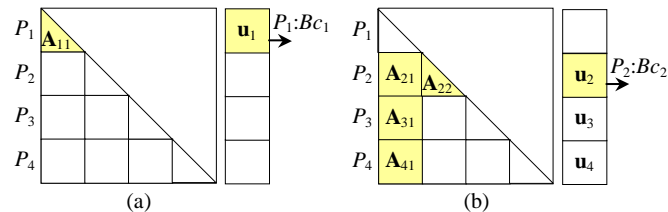
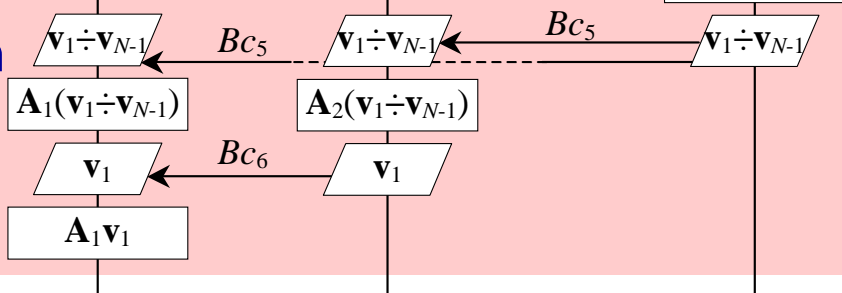
Forward calculation



Form diagonal matrix

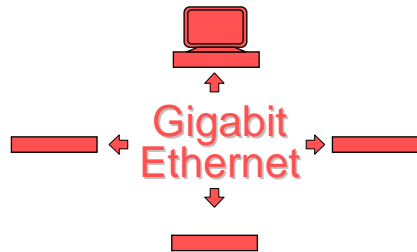


Backward substitution

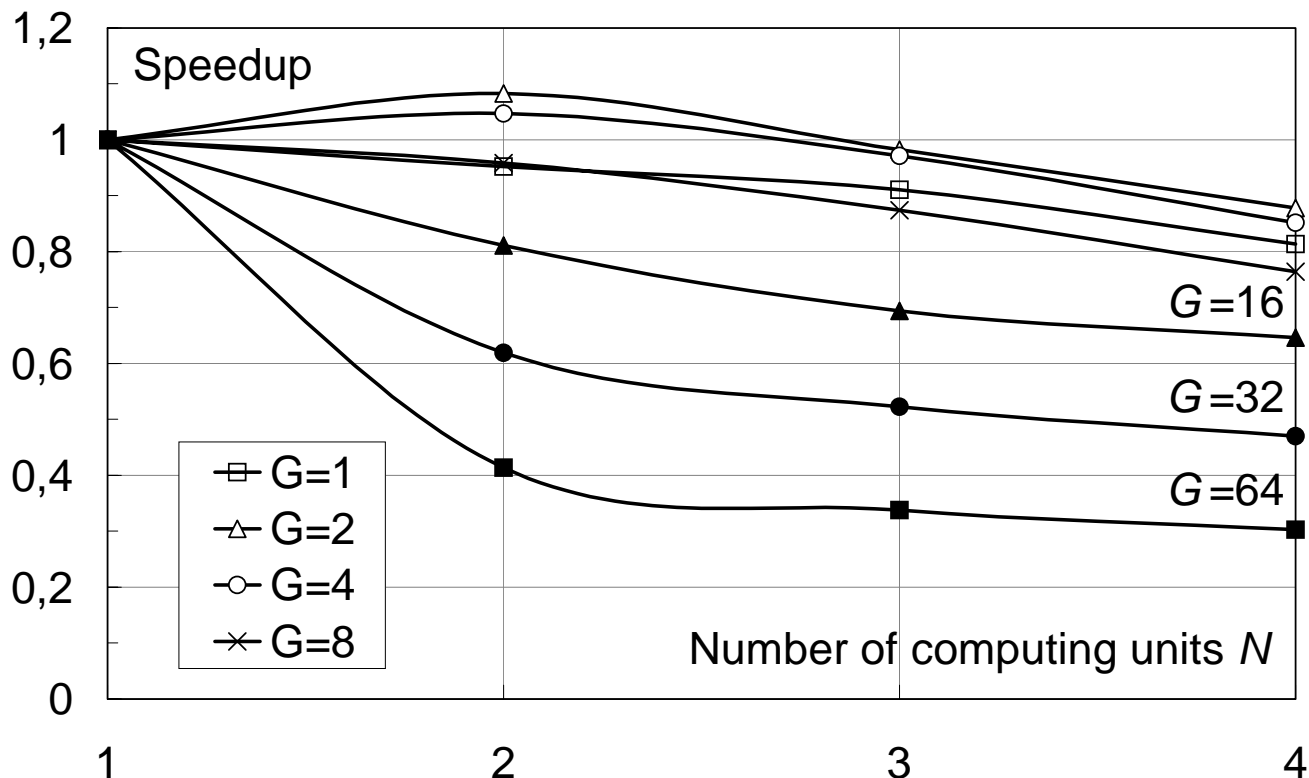


Message passing formulation

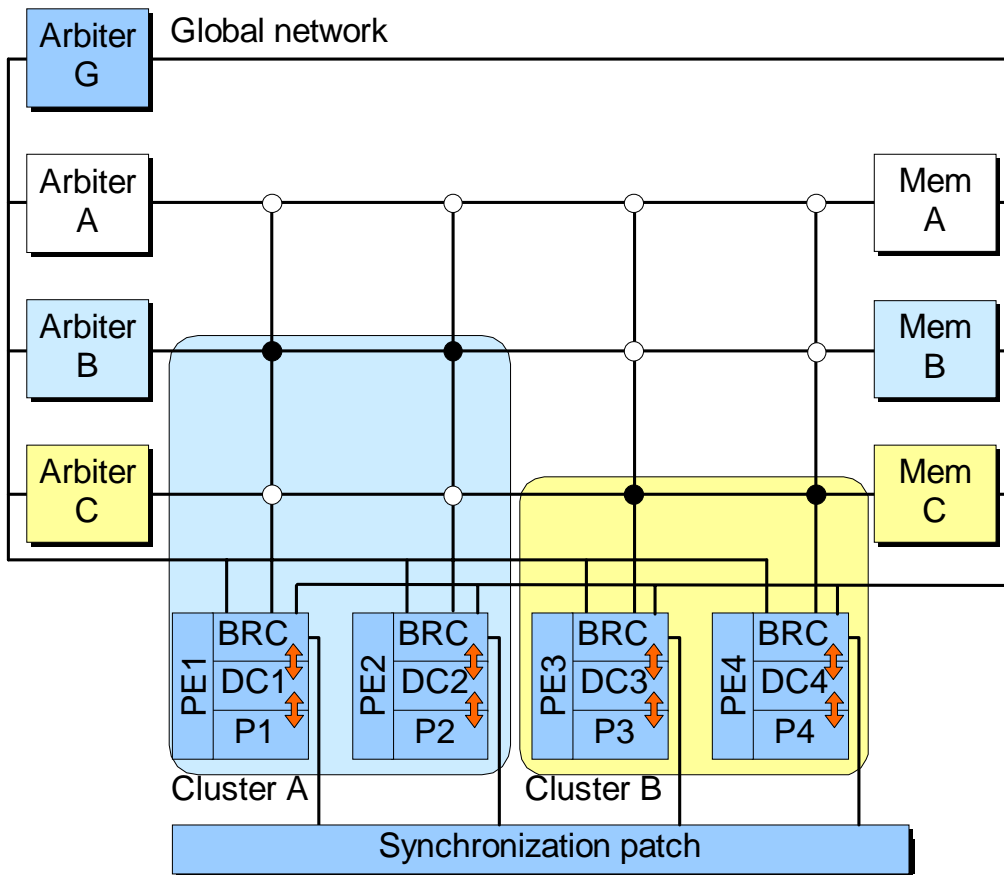
Performance analysis in a cluster of workstations:



- ◆ 4 nodes with Xeon 2,6GHz CPU,
- ◆ Gigabit-Ethernet,
- ◆ Lam-MPI library.



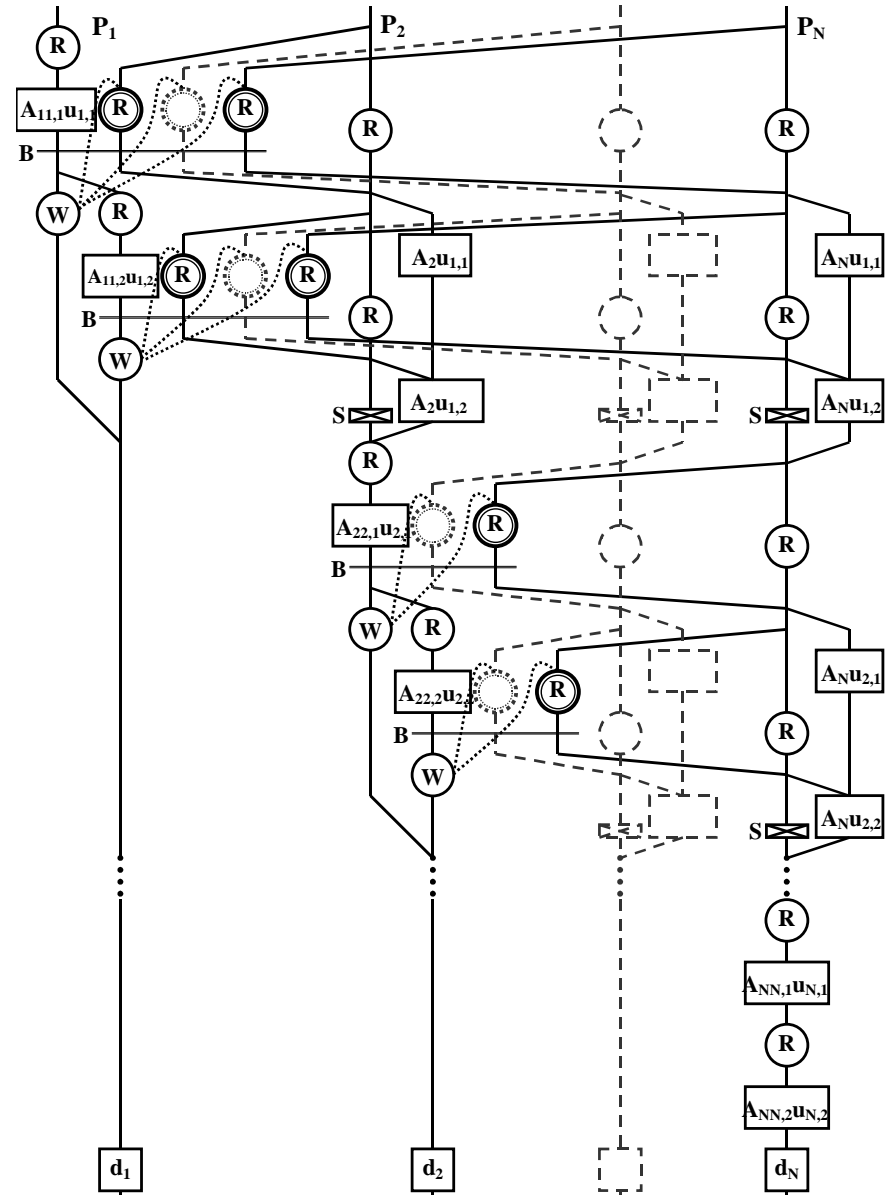
Dynamic SMP clusters



- Strategy of dynamic switching of processors between clusters based on shared memory modules and *reads on the fly* of data inside clusters at run-time (during computations).
- The permanent connection to a memory module bus is meant for communication with large data sets.
- Other processors that want to use these results, have to get connected to this memory bus dynamically, shortly before the relevant data will be sent by a producer to its permanent memory module.

Dynamic SMP clusters

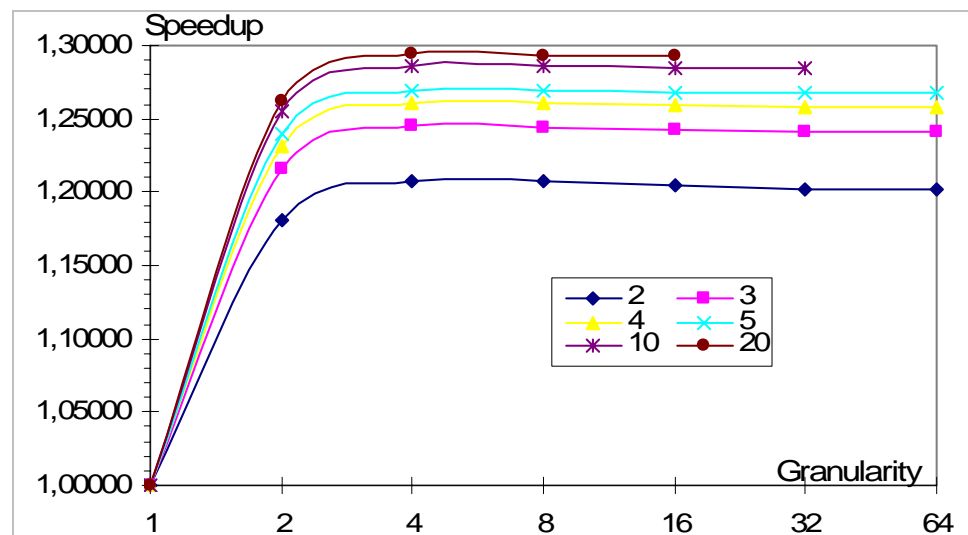
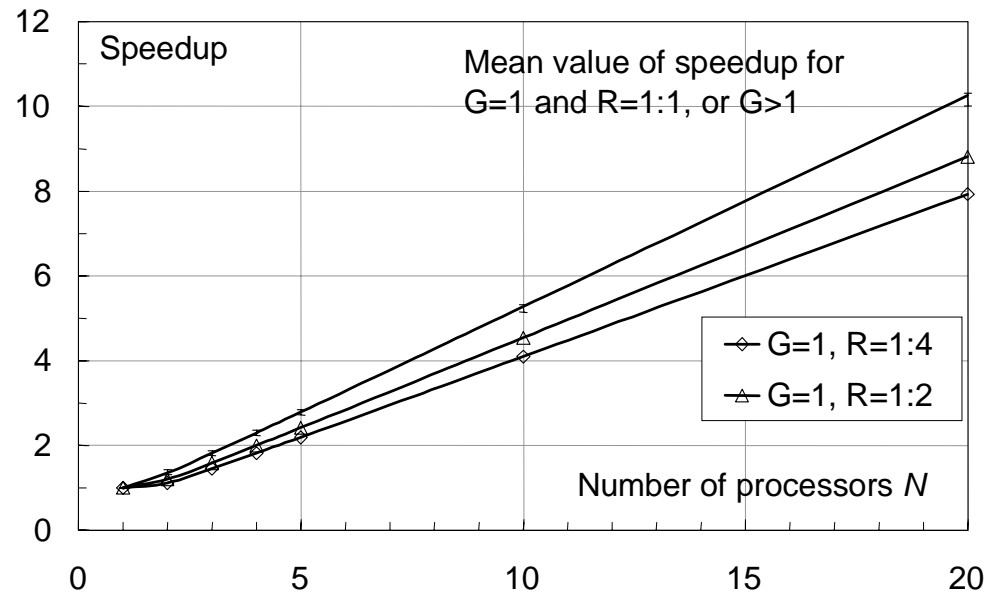
The EMDFG with reads on the fly of the SSOR preconditioner forward computations phase.



Dynamic SMP clusters

Performance analysis:

- ♦ computation speedup for SSOR in dynamic SMP clusters as a function of N (processor number) for different G and R (processor speed / memory speed).
- ♦ changes of computation speedup for SSOR in dynamic SMP clusters versus $G=1$ for $R=1:4$ as a function of G and N .



Conclusions

- ♦ Designing an efficient PCG algorithm for the multi-computer system with static architecture requires using more efficient data transmission methods.
- ♦ Two-level task decomposition enables to assure the best speedup of the iterative algorithm.
- ♦ The proposed new architecture based on dynamic SMP clusters and communication on the fly can be efficiently applied for the discussed preconditioned conjugate gradient algorithm in a parallel accelerator of the SSOR preconditioner.
- ♦ The result of the proposed task decomposition depends on the system architecture (the number of processors), system time parameters (processor and data transmission speed) and formulation of the problem.