

Calcul științific de înaltă performanță (Programarea arhitecturilor multiprocesor)

Examen de doctorat. - Programa disciplinei – Prof. Dr. Ing. Daniel Ioan

1. Arhitecturile sistemelor de calcul multiprocesor

Definiția și utilitatea procesării paralele
Arhitecturi uniprocessor - von Neumann
Arhitectura microprocesoarelor
Memoria și caracterizarea ei. Ierarhizarea.
I/O și interconectare
Arhitecturi paralele - clasificarea Flynn
Memorie: distribuită, partajată și repartizată hibrid
Consistența memoriei și modele
Memorie paralelă cu acces aleator (PRAM Parallel Random Access Memory)
Interconectarea procesoarelor
Paralelismul la nivelul CPU
Sistemelor de tip cluster și sisteme de programe specifice clusterelor de calculatoare.

2. Modele de programare paralelă

Modele de programare paralelă:
paralelismul datelor vs al taskurilor
paralelismul explicit vs implicit
memorie distribuită vs partajată
Modelele pentru memorie partajată
Modelul firelor de execuție
Modelul transmiterii mesajelor
Modelul cu paralelismul datelor
Alte modele

3. Concepția programelor paralele și distribuite

Paralelizare automată - manuală
Partitionarea problemei
Comunicația
Sincronizarea
Dependența datelor
Echilibrarea sarcinilor
Granularitatea
Intrări/ieșiri
Comunicarea sincronă și asincronă între procese
Apeluri de proceduri și metode la distanță.
Consistența și replicarea datelor: modele centrate pe date și pe client, plasarea replicilor, propagarea actualizărilor, protocoale de consistență.
Limite și costuri ale programării paralele. Complexitatea calculului paralel.
Analiza performanțelor. Legea lui Amdahl
Exemple de programare paralelă și distribuită. Algoritmi numerici. Înmulțirea matricelor. Sisteme cu matrice triunghiulare.

4. Tehnologii software pentru calcule de inalta performanta

- Message-Passing Interface (MPI)
- Parallel Virtual Machine (PVM)
- Paralelizarea automata, de compilator
- Fire paralele de executie p-thread
- OpenMP
- FORTRAN de inalta performanta (HPF)
- C++ de inalta performanta (HPC++)
- Alte tehnologii (CORBA)
- Instrumente pentru dezvoltarea si depanarea programeleor paralele
- Analiza performantelor sistemeleor paralele si distribuite
- Medii pentru rezolvarea problemelor stiintifice pe sisteme cu arhitecturi multiprocesor:
PETSc, Sca/LAPACK
- Utilizarea pachetelor MATLAB PCT si DCS
 - Cicluri paralele (parfor)
 - Modelul spmd (programe unice date multiple)
 - Calcule interactive paralele cu pmode
 - Lucrul cu tablouri codistribuite
 - Programarea cu PCT, utilizarea profiler si depanare
 - Evaluarea sincrona si asincrona a functiilor pe cluster
 - Programarea distribuita a joburilor
 - Programming Distributed Jobs
 - Programarea paralela a joburilor
- Generarea automatizată de cod paralelizat pentru arhitecturi de ultimă oră: multicore-uri omogene (Intel/AMD) și heterogene
- GPGPUs – plăci grafice (ATI/NVIDIA) -CUDA
- Soluții distribuite în sisteme heterogene bazate pe obiecte, sisteme de agenți, Grid, Cloud și sisteme P2P.
- Demonstratii si exemple de aplicatii:
 - Distribuirea tabelor
 - Rezolvarea iterativa a sistemeleor liniare cu matrice rare
 - Solutii directe pentru sisteme de mari dimensiuni cu matrice plina
 - Exemplu de utilizare PETSc,
 - Exemplu de utilizare Sca/LAPACK
 - Exemplu de utilizare MATLAB - PCT
 - Exemplu de utilizare MATLAB – DCS
 - Exemplu de utilizare MPI
 - Exemplu de utilizare OpenMP
 - Exemplu de utilizare NVIDIA-CUDA

5. Clusterul ATLAS din UPB-LMN

- Structura clusterului ATLAS
- Software comercial instlat pe ATLAS
- Software din domeniul public instalat pe ATLAS
- Resurse documentare arhivate pe ATLAS
- Demonstratii disponibile pe ATLAS
- Programe de test (benchmarks) executate pe ATLAS

Bibliografie

Pagini web introductive:

http://en.wikipedia.org/wiki/Parallel_computing OBLIGATORIU
http://en.wikipedia.org/wiki/Distributed_computing OBLIGATORIU
http://en.wikipedia.org/wiki/Grid_computing OBLIGATORIU
http://en.wikipedia.org/wiki/Automatically_Tuned_Linear_Algebra_Software
www.mathworks.com (help-ul MATLAB si articole referitoare la programarea paralela)
Inclusiv: MATLAB Distributed Computing Server 4.3: Benchmarking A\b OBLIGATORIU

Mathworks – Parallel Computing Toolbox 4- User's guide OBLIGATORIU
http://www.mathworks.com/access/helpdesk/help/pdf_doc/distcomp/distcomp.pdf

The [PVM](#) home page.

The [MPI](#) home page.

[MPICH](#) OBLIGATORIU
<http://www.mcs.anl.gov/research/projects/mpich2/>

<http://ro.wikipedia.org/wiki/CUDA>

CARTI SI RESURSE FUNDAMENTALE:

1. Ian Foster [Designing and Building Parallel Programs \(Online\)](#)
<http://www.mcs.anl.gov/~itf/dbpp/> OBLIGATORIU
2. Blaise Barney, Lawrence Livermore National Laboratory Introduction to Parallel Computing OBLIGATORIU
https://computing.llnl.gov/tutorials/parallel_comp/
3. Kumar V., Grama A., Gupta A., Karypis G. [Introduction to Parallel Computing: Design and Analysis of Algorithms](#), Ed. Benjamin-Cummings, 2003, disponibilă în format PS,
<http://www.cs.purdue.edu/homes/ayg/book/toc.html>
4. Tannenbaum A., van Steen M. [Distributed systems Principles and paradigms](#), 2002
<http://www.cs.vu.nl/~ast/books/ds1/>
5. Carti in limba romana:
 1. B. Dumitrescu Algoritmi de calcul paralel, 2001 PDF la
<http://www.schur.pub.ro/download.htm> OBLIGATORIU
 2. Cristea V., "Algoritmi de prelucrare paralelă", Ed. Matrix Rom, 2005
 3. Croitoru c., "Introducere in proiectarea algoritmilor paraleli", Ed. Matrix Rom, 2004
 4. C. Ivan , Calcul paralel si distribuit, Editura UTPress , 2007.
 5. Zaharia M.H. "Sisteme paralele și distribuite", Ed. Gh. Asachi, Iași, 2003

6. Jack Dongarra Scientific Computing for Engineers OBLIGATORIU

<http://www.cs.utk.edu/~dongarra/WEB-PAGES/cs594-2010.htm>

Aici se afla o importanta lista de resurse bibliografice din care am extras:

[The Sourcebook of Parallel Computing](#), Edited by Jack Dongarra, Ian Foster, Geoffrey Fox, William Gropp, Ken Kennedy, Linda Torczon, Andy White, October 2002, 760 pages, ISBN 1-55860-871-0, Morgan Kaufmann Publishers.

- Here's a pointer to specs on various processors:

<http://www.cpu-world.com/CPUs/index.html>

<http://www.cpu-world.com/sspec/index.html>

<http://processorfinder.intel.com>

Standard Books on Message Passing

``MPI - The Complete Reference, Volume 1, The MPI-1 Core, Second Edition",
by Marc Snir, Steve Otto, Steven Huss-Lederman, David Walker, Jack Dongarra, MIT
Press, September 1998, ISDN 0-262-69215-5.

``Using MPI,"

by William Gropp, Ewing Lusk, and Anthony Skjellum, published by MIT Press, October
1994; ISBN 0-262-57104-8.

``MPI: The Complete Reference - 2nd Edition: Volume 2 - The MPI-2 Extensions",
by William Gropp, Steven Huss-Lederman, Andrew Lumsdaine, Ewing Lusk, Bill
Nitzberg, William Saphir, and Marc Snir, published by The MIT Press, September, 1998;
ISBN 0-262-57123-4.

Other Parallel Information Sites

- [NHSE - National HPC Software Exchange](#)
- [Netlib Repository at UTK/ORNL](#)
- [BLAS Quick Reference Card](#)
- [LAPACK](#)
- [ScaLAPACK](#)
- [GAMS - Guide to Available Math Software](#)
- [Supercomputing & Parallel Computing: Conferences](#)
- [Supercomputing & Parallel Computing: Journals](#)
- [High Performance Fortran \(HPF\) reports](#)
- [High Performance Fortran Resource List](#)
- [Fortran 90 Resource List](#)
- [Major Science Research Institutions from Caltech](#)
- [Message Passing Interface \(MPI\) Forum](#)
- [High Performance Fortran Forum](#)
- [OpenMP](#)
- [PVM](#)
- [Parallel Tools Consortium](#)
- [DoD High Performance Computing Modernization Program](#)
- [DoE Accelerated Strategic Computing Initiative \(ASC\)](#)
- [National Computational Science Alliance](#)

Related On-line Textbooks

- [Templates](#) for the Solution of Linear Systems: Building Blocks for Iterative Methods, SIAM Publication, Philadelphia, 1994.
- [PVM](#) - A Users' Guide and Tutorial for Networked Parallel Computing, MIT Press, Boston, 1994.
- [MPI](#) : A Message-Passing Interface Standard
- [LAPACK](#) Users' Guide (Second Edition), SIAM Publications, Philadelphia, 1995.
- [MPI: The Complete Reference](#), MIT Press, Boston, 1996.
- [Using MPI: Portable Parallel Programming with the Message-Passing Interface](#) by W. Gropp, E. Lusk, and A. Skjellum
- [Parallel Computing Works](#), by G. Fox, R. Williams, and P. Messina (Morgan Kaufmann Publishers)
- [Computational Science Education Project](#)
- [Designing and Building Parallel Programs](#). A dead-tree version of this book is available by Addison-Wesley.
- [High Performance Fortran \(HPF\)](#), a course offered by Manchester and North High Performance Computing Training & Education Centre, United Kingdom

For performance analysis:

- Raj Jain, The Art of Computer Systems Performance Analysis. John Wiley, 1991.

Papers on performance analysis tools:

- Ruth A. Aydt, "The Pablo Self-Defining Data Format," November 1997, [click here](#).
- Jeffrey K. Hollingsworth, Barton P. Miller, Marcelo J. R. Gongalves, Oscar Naim, Zhichen Xu and Ling Zheng, "MDL: A Language and Compiler for Dynamic Program Instrumentation", International Conference on Parallel Architectures and Compilation Techniques, San Francisco, CA, November 1997, [click here](#).
- Barton P. Miller, Mark D. Callaghan, Jonathan M. Cargille, Jeffrey K. Hollingsworth, R. Bruce Irvin, Karen L. Karavanic, Krishna Kunchithapadam and Tia Newhall. "The Paradyn Parallel Performance Measurement Tools", IEEE Computer 28(11), (November 1995). [click here](#).
- Steven T. Hackstadt and Allen D. Malony, "Distributed Array Query and Visualization for High Performance Fortran", February 1996. [click here](#).
- Jerry Yan and Sekhar Sarukkai and Pankaj Mehra, "Performance Measurement, Visualization and Modeling of Parallel and Distributed Programs using the AIMS toolkit", Software Practice and Experience 25(4), April 1995, 429-461. [click here](#).

Other Online Software and Documentation

- Matlab documentation is available from several sources, most notably by typing ``help'' into the Matlab command window. A primer (for version 4.0/4.1 of Matlab, not too different from the current version) is available in either [postscript](#) or [pdf](#).
- [Netlib](#), a repository of numerical software and related documentation
- [Netlib Search Facility](#), a way to search for the software on Netlib that you need
- [GAMS - Guide to Available Math Software](#), another search facility to find numerical software
- Linear Algebra Software Libraries and Collections

- [LAPACK](#), state-of-the-art software for dense numerical linear algebra on workstations and shared-memory parallel computers. Written in Fortran.
- [CLAPACK](#), a C version of LAPACK.
(For a partial C++ version, see [LAPACK++](#) on [Roldan Pozo's homepage](#))
- [LAPACK Manual](#)
- [ScaLAPACK](#), a partial version of LAPACK for distributed-memory parallel computers.
- [ScaLAPACK manual](#)
- [LINPACK](#) and [EISPACK](#) are precursors of LAPACK, dealing with linear systems and eigenvalue problems, respectively.
- [SuperLU](#) is a fast implementations of sparse Gaussian elimination for sequential and parallel computers, respectively.
- Sources of test matrices for sparse matrix algorithms
 - [Matrix Market](#)
 - [University of Florida Sparse Matrix Collection](#)
- [Templates for the solution of linear systems](#), a collection of iterative methods, with advice on which ones to use. The web site includes on-line versions of the book (in [html](#) and [postscript](#)) as well as software.
- [Templates for the Solution of Algebraic Eigenvalue Problems](#) is a survey of algorithms and software for solving eigenvalue problems. The web site points to an html version of the book, as well as software.
 - [Updated survey](#) of sparse direct linear equation solvers, by [Xiaoye Li](#)
- [MGNet](#) is a repository for information and software for Multigrid and Domain Decomposition methods, which are widely used methods for solving linear systems arising from PDEs.
- Resources for Parallel and High Performance Computing
 - [Millennium](#) a UC Berkeley campus-wide parallel computing resource
 - [Resources for CS 267, Applications of Parallel Computers](#)
 - [ACTS \(Advanced CompuTational Software\)](#) is a set of software tools that make it easier for programmers to write high performance scientific applications for parallel computers.
 - [PETSc: Portable, Extensible, Toolkit for Scientific Computation](#)
 - [NHSE - National High Performance Computing and Communications Software Exchange](#), pointers to related work across the country.
- Issues related to Computer Arithmetic and Error Analysis
 - Efficient software for [very high precision floating point arithmetic](#)
 - [Notes on IEEE Floating Point Arithmetic](#), by [Prof. W. Kahan](#)
 - [Other notes on arithmetic, error analysis, etc.](#) by [Prof. W. Kahan](#)
 - [Report on arithmetic error that cause the Ariane 5 Rocket Crash](#)
 - The IEEE floating point standard is currently being updated. To find out what issues the standard committee is considering, look [here](#).