

Lecture Notes in Computer Science
Edited by G. Goos, J. Hartmanis, and J. van Leeuwen

2840

Springer

Berlin

Heidelberg

New York

Hong Kong

London

Milan

Paris

Tokyo

Jack Dongarra Domenico Laforenza
Salvatore Orlando (Eds.)

Recent Advances in Parallel Virtual Machine and Message Passing Interface

10th European PVM/MPI User's Group Meeting
Venice, Italy, September 29 - October 2, 2003
Proceedings



Springer

Series Editors

Gerhard Goos, Karlsruhe University, Germany
Juris Hartmanis, Cornell University, NY, USA
Jan van Leeuwen, Utrecht University, The Netherlands

Volume Editors

Jack Dongarra
University of Tennessee, Computer Science Department
1122 Volunteer Blvd, Knoxville, TN 37996-3450, USA
E-mail: dongarra@cs.utk.edu

Domenico Laforenza
Istituto di Scienza e Technologie dell'Informazione
Consiglio Nazionale delle Ricerche (ISTI-CNR)
Area della Ricerca CNR, Via G. Moruzzi, 1, 56126 Pisa, Italy
E-mail: domenico.laforenza@isti.cnr.it

Salvatore Orlando
Università Ca' Foscari di Venezia, Dipartimento di Informatica
Via Torino, 155, 30172 Venezia Mestre, Italy
E-mail: orlando@dsi.unive.it

Cataloging-in-Publication Data applied for

A catalog record for this book is available from the Library of Congress.

Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie;
detailed bibliographic data is available in the Internet at <<http://dnb.ddb.de>>.

CR Subject Classification (1998): D.1.3, D.3.2, F.1.2, G.1.0, B.2.1, C.1.2

ISSN 0302-9743

ISBN 3-540-20149-1 Springer-Verlag Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable for prosecution under the German Copyright Law.

Springer-Verlag Berlin Heidelberg New York
a member of BertelsmannSpringer Science+Business Media GmbH

<http://www.springer.de>

© Springer-Verlag Berlin Heidelberg 2003
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Olgun Computergrafik
Printed on acid-free paper SPIN: 10961599 06/3142 5 4 3 2 1 0

Preface

The message passing paradigm is considered the most effective way to develop efficient parallel applications. PVM (Parallel Virtual Machine) and MPI (Message Passing Interface) are the most frequently used tools for programming message passing applications.

This volume includes the selected contributions presented at the 10th European PVM/MPI Users' Group Meeting (Euro PVM/MPI 2003), which was held in Venice, Italy, September 29–October 2, 2003. The conference was jointly organized by the Department of Computer Science of the Ca' Foscari University of Venice, Italy and the Information Science and Technologies Institute of the National Research Council (ISTI-CNR), Pisa, Italy.

The conference was previously held in Linz, Austria (2002), Santorini, Greece (2001), Balatonfüred, Hungary (2000), Barcelona, Spain (1999), Liverpool, UK (1998), and Krakow, Poland (1997). The first three conferences were devoted to PVM and were held in Munich, Germany (1996), Lyon, France (1995), and Rome, Italy (1994).

The conference has become a forum for users and developers of PVM, MPI, and other message passing environments. Interactions between these groups has proved to be very useful for developing new ideas in parallel computing, and for applying some of those already existent to new practical fields. The main topics of the meeting were evaluation and performance of PVM and MPI, extensions, implementations and improvements of PVM and MPI, parallel algorithms using the message passing paradigm, and parallel applications in science and engineering. In addition, the topics of the conference were extended to include Grid computing, in order to reflect the importance of this area for the high-performance computing community.

This year we received 115 submissions, and the Program Committee finally selected 64 regular papers, and 16 short papers. Besides the main track of contributed papers, the conference featured the second edition of the special session "ParSim 03 – Current Trends in Numerical Simulation for Parallel Engineering Environments." This volume also includes six short papers presented during the ParSim 03 session.

Two tutorials were presented during the meeting: "High-Level Programming in MPI" by William Gropp and Ewing Lusk, and "Programming Environments for Grids and Distributed Computing Systems" by Vaidy Sunderam. Finally, six invited talks were presented at the conference: the invited speakers were Geoffrey Fox, Al Geist, William Gropp, Ewing Lusk, Thierry Priol, and Marco Vanneschi. The contributions relating to the invited talks and tutorials are also included in this volume.

We would like to express our gratitude for the kind support of our sponsors (see below). We are also indebted to all the members of the Program Committee and the additional reviewers, who ensured the high quality of Euro PVM/MPI

2003 with their careful and timely work. Finally, we would like to express our gratitude to our colleagues in the ISTI-CNR and the University of Venice for their help and support during the conference organization. In particular, we would like to thank Tiziano Fagni, Paolo Palmerini, Raffaele Perego, Claudio Silvestri, Fabrizio Silvestri and Nicola Tonellotto.

September 2003

Jack Dongarra
Domenico Laforenza
Salvatore Orlando

VIII Program Committee

Neil D. Pundit	Sandia National Labs, USA
Rolf Rabenseifner	University of Stuttgart, Germany
Andrew Rau-Chaplin	Dalhousie University, Canada
Ralf Reussner	DSTC, Monash University, Australia
Yves Robert	ENS Lyon, France
Casiano Rodriguez-Leon	Universidad de La Laguna, Spain
Miquel Senar	Universitat Autònoma de Barcelona, Spain
Joao Gabriel Silva	University of Coimbra, Portugal
Vaidy Sunderam	Emroy University, USA
Francisco Tirado	Universidad Complutense de Madrid, Spain
Bernard Tourancheau	SUN Labs Europe, France
Jesper Larsson Träff	NEC Europe Ltd., Germany
Pavel Tvrdik	Czech Technical University, Czech Republic
Umberto Villano	Università del Sannio, Italy
Jens Volkert	Joh. Kepler University, Linz, Austria
Jerzy Wasniewski	Danish Computing Centre, Denmark
Roland Wismüller	Technical Univ. of Munich, Germany

Additional Reviewers

Alba, Enrique
Almeida, Francisco
Angskun, Thara
Balis, Bartosz
Balogh, Zoltan
Baltzer, Oliver
Bautista, Alfredo
Beyls, Kristof
Bönisch, Homas
Bosilca, George
Brandes, Thomas
Caron, Eddy
Cortes, Ana
Cronk, David
De Sande, Francisco
Dobruchy, Miroslav
Eavis, Todd
Fagni, Tiziano
Ferrini, Renato
Franco, Daniel
Funika, Wlodzimierz
Fürlinger, Karl
Gabriel, Edgar
Garcia, Carlos
Gelas, Jean-Patrick
Giang, Nguyen T.
Guérin-Lassous, Isabelle
Habala, Ondrej
Halada, Ladislav
Haubold, Sven
Hermann, Gerd
Hernández, Porfidio
Hetze, Bernd
Heymann, Elisa
Iacono, Mauro
Keller, Ainer
Krämer-Fuhrmann, Ottmar
Legrand, Arnaud
León, Coromoto
L'Excellent, Jean-Yves
Luksch, Eter
Mairandres, Martin
Malawski, Maciej
Mancini, Emilio P.
Margalef, Tomas
Marques, Rui
Medeiros, Pedro
Mix, Hartmut
Moscato, Francesco
Müller-Pfefferkorn, Ralph
Nemeth, Zsolt
Palmerini, Paolo
Pedretti, Kevin
Pflüger, Stefan
Philippe, Laurent
Prieto-Matias, Manuel
Puppini, Diego
Rak, Massimiliano
Rastello, Fabrice
Rathmayer, Sabine
Renard, Hélène
Ripoll, Ana
Schaubschlaeger, Christian
Schmidt, Andreas
Senar, Miquel A.
Silvestri, Fabrizio
Simo, Branislav
Stamatakis, Alexandros
Suppi, Remo
Tonello, Nicola
Tran, Viet D.
Underwood, Keith
Venticinque, Salvatore
Vivien, Frédéric
Walter, Max
Weidendorfer, Osef
Worringen, Joachim
Zajac, Katarzyna

Sponsoring Institutions

(as of July 18th, 2002)

HP (Hewlett-Packard)

Sun Microsystems

Microsoft

ONRIFO (US Office of Naval Research – International Field Office)

Critical Software

DATAMAT SpA

Department of Computer Science, Ca' Foscari University of Venice

Information Science and Technologies Institute, National Research Council
(ISTI-CNR), Pisa

Table of Contents

Invited Talks

Messaging Systems: Parallel Computing the Internet and the Grid	1
<i>G. Fox</i>	
Progress towards Petascale Virtual Machines	10
<i>A. Geist</i>	
Future Developments in MPI	15
<i>W.D. Gropp</i>	
Integrating Scalable Process Management into Component-Based Systems Software	16
<i>E. Lusk</i>	
Programming High Performance Applications Using Components	23
<i>T. Priol</i>	
ASSIST High-Performance Programming Environment: Application Experiences and Grid Evolution	24
<i>M. Vanneschi</i>	

Tutorials

High-Level Programming in MPI	27
<i>W. Gropp and E. Lusk</i>	
Programming Environments for Grids and Distributed Computing Systems	28
<i>V. Sunderam</i>	

Evaluation and Performance Analysis

Performance Modeling and Evaluation of Java Message-Passing Primitives on a Cluster	29
<i>G.L. Taboada, J. Touriño, and R. Doallo</i>	
Integrating New Capabilities into NetPIPE	37
<i>D. Turner, A. Oline, X. Chen, and T. Benjergedes</i>	
Off-Line Performance Prediction of Message-Passing Applications on Cluster Systems	45
<i>E. Mancini, M. Rak, R. Torella, and U. Villano</i>	

Complexity Driven Performance Analysis	55
<i>L. García, J.A. González, J.C. González, C. León, C. Rodríguez, and G. Rodríguez</i>	
Usefulness and Usage of SKaMPI-Bench	63
<i>W. Augustin and T. Worsch</i>	
The Performance of Parallel Disk Write Methods for Linux Multiprocessor Nodes	71
<i>G.D. Benson, K. Long, and P.S. Pacheco</i>	
A Model for Performance Analysis of MPI Applications on Terascale Systems	81
<i>S. Chakravarthi, C.R. Krishna Kumar, A. Skjellum, H.A. Prahalad, and B. Seshadri</i>	
Evaluating the Performance of MPI-2 Dynamic Communicators and One-Sided Communication	88
<i>E. Gabriel, G.E. Fagg, and J.J. Dongarra</i>	
Ring Algorithms on Heterogeneous Clusters with PVM: Performance Analysis and Modeling	98
<i>A. Corana</i>	
An MPI Tool to Measure Application Sensitivity to Variation in Communication Parameters	108
<i>E.A. León, A.B. Maccabe, and R. Brightwell</i>	
Measuring MPI Latency Variance	112
<i>R. Riesen, R. Brightwell, and A.B. Maccabe</i>	
Parallel Algorithms Using Message Passing	
CGMgraph/CGMlib: Implementing and Testing CGM Graph Algorithms on PC Clusters	117
<i>A. Chan and F. Dehne</i>	
Efficient Parallel Implementation of Transitive Closure of Digraphs	126
<i>C.E.R. Alves, E.N. Cáceres, A.A. Castro Jr, S.W. Song, and J.L. Szwarcfiter</i>	
A Scalable Crystallographic FFT	134
<i>J. Sequel and D. Burbano</i>	
Object-Oriented NeuroSys: Parallel Programs for Simulating Large Networks of Biologically Accurate Neurons	142
<i>P.S. Pacheco, P. Miller, J. Kim, T. Leese, and Y. Zabiyyaka</i>	
PageRank Computation Using PC Cluster	152
<i>A. Rungswang and B. Manaskasemsak</i>	

An Online Parallel Algorithm for Remote Visualization of Isosurfaces	160
<i>A. Clematis, D. D'Agostino, and V. Gianuzzi</i>	
Parallel Algorithms for Computing the Smith Normal Form of Large Matrices	170
<i>G. Jäger</i>	
Hierarchical MPI+OpenMP Implementation of Parallel PIC Applications on Clusters of Symmetric MultiProcessors	180
<i>S. Briguglio, B. Di Martino, G. Fogaccia, and G. Vlad</i>	
Non-strict Evaluation of the FFT Algorithm in Distributed Memory Systems	188
<i>A. Cristóbal-Salas, A. Tchernykh, and J.-L. Gaudiot</i>	
A Parallel Approach for the Solution of Non-Markovian Petri Nets	196
<i>M. Scarpa, S. Distefano, and A. Puliafito</i>	
Advanced Hybrid MPI/OpenMP Parallelization Paradigms for Nested Loop Algorithms onto Clusters of SMPs	204
<i>N. Drosinos and N. Koziris</i>	
The AGEB Algorithm for Solving the Heat Equation in Two Space Dimensions and Its Parallelization on a Distributed Memory Machine	214
<i>N. Alias, M.S. Sahimi, and A.R. Abdullah</i>	
A Parallel Scheme for Solving a Tridiagonal Matrix with Pre-propagation	222
<i>A. Wakatani</i>	
Competitive Semantic Tree Theorem Prover with Resolutions	227
<i>C.K. Kim and M. Newborn</i>	
Explicit Group Iterative Solver on a Message Passing Environment	232
<i>M.A. Norhashidah Hj., A. Rosni, and J.L. Kok</i>	
Applying Load Balancing in Data Parallel Applications Using DASUD	237
<i>A. Cortés, M. Planas, J.L. Millán, A. Ripoll, M.A. Senar, and E. Luque</i>	
Performance Analysis of Approximate String Searching Implementations for Heterogeneous Computing Platform	242
<i>P.D. Michailidis and K.G. Margaritis</i>	

Extensions, Improvements and Implementations of PVM/MPI

Using a Self-connected Gigabit Ethernet Adapter as a memcpy() Low-Overhead Engine for MPI	247
<i>G. Ciaccio</i>	

Improving the Performance of Collective Operations in MPICH	257
<i>R. Thakur and W.D. Gropp</i>	
PVMWebCluster: Integration of PVM Clusters Using Web Services and CORBA	268
<i>P. Czarnul</i>	
Lock-Free Collective Operations	276
<i>A. Supalov</i>	
Efficient Message-Passing within SMP Systems	286
<i>X. Chen and D. Turner</i>	
The Network Agnostic MPI – Scali MPI Connect	294
<i>L.P. Huse and O.W. Saastad</i>	
PC/MPI: Design and Implementation of a Portable MPI Checkpointer . . .	302
<i>S. Ahn, J. Kim, and S. Han</i>	
Improving Generic Non-contiguous File Access for MPI-IO	309
<i>J. Worringen, J. Larsson Trüff, and H. Ritzdorf</i>	
Remote Exception Handling for PVM Processes	319
<i>P.L. Kaczmarek and H. Krawczyk</i>	
Evaluation of an Eager Protocol Optimization for MPI	327
<i>R. Brightwell and K. Underwood</i>	
A Comparison of MPICH Allgather Algorithms on Switched Networks . . .	335
<i>G.D. Benson, C.-W. Chu, Q. Huang, and S.G. Caglar</i>	
Network Fault Tolerance in LA-MPI	344
<i>R.T. Aulwes, D.J. Daniel, N.N. Desai, R.L. Graham, L.D. Risinger, M.W. Sukalski, and M.A. Taylor</i>	
MPI on BlueGene/L: Designing an Efficient General Purpose Messaging Solution for a Large Cellular System	352
<i>G. Almási, C. Archer, J.G. Castaños, M. Gupta, X. Martorell, J.E. Moreira, W.D. Gropp, S. Rus, and B. Toonen</i>	
Porting P4 to Digital Signal Processing Platforms	362
<i>J.A. Rico, J.C. Díaz Martín, J.M. Rodríguez García, J.M. Álvarez Llorente, and J.L. García Zapata</i>	
Fast and Scalable Barrier Using RDMA and Multicast Mechanisms for InfiniBand-Based Clusters	369
<i>S.P. Kini, J. Liu, J. Wu, P. Wyckoff, and D.K. Panda</i>	
A Component Architecture for LAM/MPI	379
<i>J.M. Squyres and A. Lumsdaine</i>	

ORNL-RSH Package and Windows '03 PVM 3.4	388
<i>P. Pfeiffer, S.L. Scott, and H. Shukla</i>	
MPI for the Clint Gb/s Interconnect	395
<i>N. Fugier, M. Herbert, E. Lemoine, and B. Tourancheau</i>	
Implementing Fast and Reusable Datatype Processing	404
<i>R. Ross, N. Miller, and W.D. Gropp</i>	
An MPI Implementation Supported by Process Migration and Load Balancing	414
<i>A. Maloney, A. Goscinski, and M. Hobbs</i>	
PVM over the CLAN Network	424
<i>R. Sohan and S. Pope</i>	

Parallel Programming Tools

Distributed Configurable Application Monitoring on SMP Clusters	429
<i>K. Frlinger and M. Gerndt</i>	
Integrating Multiple Implementations and Structure Exploitation in the Component-Based Design of Parallel ODE Solvers	438
<i>J.M. Mantas, J. Ortega Lopera, and J.A. Carrillo</i>	
Architecture of Monitoring System for Distributed Java Applications	447
<i>M. Bubak, W. Funika, M. Smętek, Z. Kiliański, and R. Wismller</i>	
A Communication API for Implementing Irregular Algorithms on SMP Clusters	455
<i>J. Hippold and G. Rnger</i>	
TOM – Efficient Monitoring Infrastructure for Multithreaded Programs ...	464
<i>B. Bali, M. Bubak, W. Funika, R. Wismller, and G. Kaplita</i>	
MPI Farm Programs on Non-dedicated Clusters	473
<i>N. Fonseca and J.G. Silva</i>	
Application Composition in Ensemble Using Intercommunicators and Process Topologies	482
<i>Y. Cotronis</i>	
Improving Properties of a Parallel Program in ParJava Environment	491
<i>V. Ivannikov, S. Gaissaryan, A. Avetisyan, and V. Padaryan</i>	

Applications in Science and Engineering

Flow Pattern and Heat Transfer Rate in Three-Dimensional Rayleigh-Benard Convection	495
<i>T. Watanabe</i>	

A Parallel Split Operator Method for the Time Dependent Schrödinger Equation	503
<i>J.P. Hansen, T. Matthey, and T. Sørveik</i>	
A Parallel Software for the Reconstruction of Dynamic MRI Sequences . . .	511
<i>G. Landi, E. Loli Piccolomini, and F. Zama</i>	
Improving Wildland Fire Prediction on MPI Clusters	520
<i>B. Abdalhaq, G. Bianchini, A. Cortés, T. Margalef, and E. Luque</i>	
Building 3D State Spaces of Virtual Environments with a TDS-Based Algorithm	529
<i>A. Křenek, I. Peterlík, and L. Matyska</i>	
Parallel Pencil-Beam Redefinition Algorithm	537
<i>P. Alderson, M. Wright, A. Jain, and R. Boyd</i>	
Dynamic Load Balancing for the Parallel Simulation of Cavitating Flows	545
<i>F. Wrona, P.A. Adamidis, U. Iben, R. Rabenseifner, and C.-D. Munz</i>	
Message Passing Fluids: Molecules as Processes in Parallel Computational Fluids	550
<i>G. Argentini</i>	
Parallel Implementation of Interval Analysis for Equations Solving	555
<i>Y. Papegay, D. Daney, and J.-P. Merlet</i>	
A Parallel System for Performing Colonic Tissue Classification by Means of a Genetic Algorithm	560
<i>S.A. Amin, J. Filippas, R.N.G. Naquib, and M.K. Bennett</i>	
Eigenanalysis of Finite Element 3D Flow Models by Parallel Jacobi–Davidson	565
<i>L. Bergamaschi, A. Martinez, G. Pini, and F. Sartoretto</i>	
Grid and Heterogeneous Computing	
Executing and Monitoring PVM Programs in Computational Grids with Jini	570
<i>G. Sipos and P. Kacsuk</i>	
Multiprogramming Level of PVM Jobs in a Non-dedicated Linux NOW . . .	577
<i>F. Giné, F. Solsona, J. Barrientos, P. Hernández, M. Hanzich, and E. Luque</i>	
Mapping and Load-Balancing Iterative Computations on Heterogeneous Clusters	586
<i>A. Legrand, H. Renard, Y. Robert, and F. Vivien</i>	

Dynamic Topology Selection for High Performance MPI in the Grid Environments	595
<i>K.-L. Park, H.-J. Lee, K.-W. Koh, O.-Y. Kwon, S.-Y. Park, H.-W. Park, and S.-D. Kim</i>	
Monitoring Message Passing Applications in the Grid with GRM and R-GMA	603
<i>N. Podhorszki and P. Kacsuk</i>	
Component-Based System for Grid Application Workflow Composition . . .	611
<i>M. Bubak, K. Górká, T. Gubała, M. Malawski, and K. Zajac</i>	
Evaluating and Enhancing the Use of the GridFTP Protocol for Efficient Data Transfer on the Grid	619
<i>M. Cannataro, C. Mastroianni, D. Talia, and P. Trunfio</i>	
Resource Monitoring and Management in Metacomputing Environments . .	629
<i>T. Wrzosek, D. Kurzyniec, D. Drzewiecki, and V. Sunderam</i>	
Generating an Efficient Dynamics Multicast Tree under Grid Environment	636
<i>T. Vorakosit and P. Uthayopas</i>	
Topology-Aware Communication in Wide-Area Message-Passing	644
<i>C.A. Lee</i>	
Design and Implementation of Dynamic Process Management for Grid-Enabled MPICH	653
<i>S. Kim, N. Woo, H.Y. Yeom, T. Park, and H.-W. Park</i>	
Scheduling Tasks Sharing Files on Heterogeneous Clusters	657
<i>A. Giersch, Y. Robert, and F. Vivien</i>	
Special Session: ParSim 03	
Special Session of EuroPVM/MPI 2003: Current Trends in Numerical Simulation for Parallel Engineering Environments – ParSim 2003	661
<i>C. Trinitis and M. Schulz</i>	
Efficient and Easy Parallel Implementation of Large Numerical Simulations	663
<i>R. Revire, F. Zara, and T. Gautier</i>	
Toward a Scalable Algorithm for Distributed Computing of Air-Quality Problems	667
<i>M. Garbey, R. Keller, and M. Resch</i>	
A Piloting SIMulator for Maritime and Fluvial Navigation: SimNav	672
<i>M. Vayssade and A. Pourplanche</i>	

Methods and Experiences of Parallelizing Flood Models 677
L. Hluchy, V.D. Tran, D. Froehlich, and W. Castaings

padfem2 – An Efficient, Comfortable Framework
for Massively Parallel FEM-Applications 681
S. Blazy, O. Kao, and O. Marquardt

AUTOBENCH/AUTO-OPT: Towards an Integrated Construction
Environment for Virtual Prototyping in the Automotive Industry 686
A. Kuhlmann, C.-A. Thole, and U. Trottenberg

Author Index 691