

Franck Cappello  
Thomas Herault  
Jack Dongarra (Eds.)

LNCS 4757

# Recent Advances in Parallel Virtual Machine and Message Passing Interface

14th European PVM/MPI Users' Group Meeting  
Paris, France, September/October 2007  
Proceedings

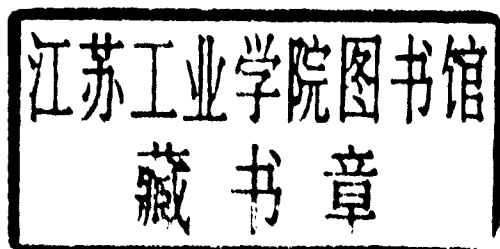


Springer

Franck Cappello Thomas Herault  
Jack Dongarra (Eds.)

# Recent Advances in Parallel Virtual Machine and Message Passing Interface

14th European PVM/MPI Users' Group Meeting  
Paris, France, September 30 - October 3, 2007  
Proceedings



## Volume Editors

Franck Cappello  
INRIA Futurs  
LRI, Bat 490 University Paris South  
91405 Orsay France  
E-mail: franck.cappello@lri.fr

Thomas Herault  
Universite Paris Sud-XI  
Laboratoire de Recherche en Informatique  
Bâtiment 490  
91405 Orsay France  
E-mail: thomas.herault@lri.fr

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

Library of Congress Control Number: 2007935601

CR Subject Classification (1998): D.1.3, D.2.3, F.1.2, G.1.0, B.2.1, B.2.1, C.1.2

LNCS Sublibrary: SL 2 – Programming and Software Engineering

|         |                                                       |
|---------|-------------------------------------------------------|
| ISSN    | 0302-9743                                             |
| ISBN-10 | 3-540-75415-6 Springer Berlin Heidelberg New York     |
| ISBN-13 | 978-3-540-75415-2 Springer 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. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media

[springer.com](http://springer.com)

© Springer-Verlag Berlin Heidelberg 2007  
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India  
Printed on acid-free paper      SPIN: 12168712      06/3180      5 4 3 2 1 0

*Commenced Publication in 1973*

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

## Editorial Board

David Hutchison

*Lancaster University, UK*

Takeo Kanade

*Carnegie Mellon University, Pittsburgh, PA, USA*

Josef Kittler

*University of Surrey, Guildford, UK*

Jon M. Kleinberg

*Cornell University, Ithaca, NY, USA*

Friedemann Mattern

*ETH Zurich, Switzerland*

John C. Mitchell

*Stanford University, CA, USA*

Moni Naor

*Weizmann Institute of Science, Rehovot, Israel*

Oscar Nierstrasz

*University of Bern, Switzerland*

C. Pandu Rangan

*Indian Institute of Technology, Madras, India*

Bernhard Steffen

*University of Dortmund, Germany*

Madhu Sudan

*Massachusetts Institute of Technology, MA, USA*

Demetri Terzopoulos

*University of California, Los Angeles, CA, USA*

Doug Tygar

*University of California, Berkeley, CA, USA*

Moshe Y. Vardi

*Rice University, Houston, TX, USA*

Gerhard Weikum

*Max-Planck Institute of Computer Science, Saarbruecken, Germany*

# Preface

Thirteen years after the publication of the first MPI (message passing interface) specification and 17 years after the first published paper on PVM (parallel virtual machine), MPI and PVM have emerged as standard programming environments and continue to be the development environment of choice for a large variety of applications, hardware platforms, and usage scenarios. There are many reasons behind this success and one of them is certainly the strength of its community.

EuroPVM/MPI is a flagship conference for this community, established as the premier international forum for researchers, users, and vendors to present their latest advances in MPI and PVM. EuroPVM/MPI is the forum where fundamental aspects of message passing, implementations, standards, benchmarking, performance, and new techniques are presented and discussed by researchers, developers and users from academia and industry.

EuroPVM/MPI 2007 was organized by INRIA in Paris, September 29 to October 3, 2007. This was the 14th issue of the conference, which takes place each year at a different European location. Previous meetings were held in Bonn (2006), Sorrento (2005), Budapest (2004), Venice (2003), Linz (2002), Santorini (2001), Balatonfüred (2000), Barcelona (1999), Liverpool (1998), Krakow (1997), Munich (1996), Lyon (1995), and Rome (1994).

The main topics of the meeting were collective operations, one-sided communication, parallel applications using the message passing paradigm, MPI standard extensions or evolution, fault tolerance, formal verification of MPI processes, MPI-I/O, performance evaluation, and hierarchical infrastructures (parallel computers and Grids).

For this year's conference, the Program Committee Co-chairs invited six outstanding researchers to present lectures on different aspects of the message passing paradigm: Tony Hey, who co-authored the first draft for the MPI standard, presented "MPI: Past, Present and Future," Al Geist, one of the authors of PVM, presented "Sustained PetaScale, the Next MPI Challenge," Satoshi Matsuoka, a pioneer of the Grid Computing, presented "The Tsubame Cluster Experience," Ewing Lusk, one of the leaders of MPICH, presented "New and Old Tools and Programming Models for High-Performance Computing," George Bosilca, one of the leading members of OpenMPI, presented "The X-Scale Challenge," and Bernd Mohr, a pioneer in performance analysis tools for Parallel Computing, presented "To Infinity and Beyond!?"

In addition to the conference main track, the meeting featured the sixth edition of the special session "ParSim 2007 – Current Trends in Numerical Simulation for Parallel Engineering Environments." The conference also included three tutorials, one on "Using MPI-2: A Problem-Based Approach" by William Gropp and Ewing Lusk, the second by Stephen Siegel on "Verifying Parallel Programs with MPI-Spin," and the third by George Bosilca and Julien Langou on "Advanced MPI Programming."

Contributions to EuroPVM/MPI 2007 and the special session ParSim were submitted in May and June, respectively. Out of the 68 submitted full papers, 40 were selected for presentation at the conference. The task of reviewing was carried out smoothly within very strict time limits by a large Program Committee, which included members from most of the American and European groups involved in MPI and PVM development, as well as from significant user communities. Almost all papers received four reviews, some even five, and none fewer than three, which provided a solid basis for the Program Chairs to make the final selection for the conference program. The result was a well-balanced, focused, and high-quality program. Out of the accepted 40 papers, four were selected as outstanding contributions to EuroPVM/MPI 2007, and were presented in special, plenary sessions:

- “Full Bandwidth Broadcast, Reduction and Scan With Only Two Trees” by Peter Sanders, Jochen Speck and Jesper Larsson Traff
- “Process Cooperation in Multiple Message Broadcast” by Bin Jia
- “Self-Consistent MPI Performance Requirements” by Jesper Larsson Traff, William Gropp and Rajeev Thakur
- “Test Suite for Evaluating Performance of MPI Implementations That Support MPI\_THREAD\_MULTIPLE” by Rajeev Thakur and William Gropp

An important part of EuroPVM/MPI is the technically oriented vendor session. At EuroPVM/MPI 2007 seven significant vendors of hardware and software for high-performance computing (Microsoft, Hewlett Packard, IBM, CISCO, Myricom, Intel, and Voltaire) presented their latest products and developments.

Information about the conference can be found at the conference Web site: <http://pvmmpi07.lri.fr/>, which will be kept available.

The proceedings were edited by Franck Cappello and Thomas Herault. The EuroPVM/MPI 2007 logo was designed by Ala Rezmerita.

The program and general chairs would like to thank all who contributed to making EuroPVM/MPI 2007 a fruitful and stimulating meeting, be they technical paper or poster authors, Program Committee members, external referees, participants, or sponsors. We would like to express our gratitude to all the members of the Program Committee and the additional reviewers, who ensured the high quality of Euro PVM/MPI 2007 with their careful work.

Finally, we would like to thank deeply INRIA and LRI for their support and efforts in organizing this event. In particular, we would like to thank Catherine Girard (INRIA), Nicole Lefevre (LRI), Gaelle Dorkeld (INRIA), and Chantal Girodon (INRIA). A special thanks to all PhD students who helped in the logistics of the conference.

October 2007

Franck Cappello  
Thomas Herault  
Jack Dongarra



# Organization

## General Chair

Jack J. Dongarra

University of Tennessee, Knoxville, USA

## Program Chairs

Franck Cappello

INRIA

Thomas Herault

Université Paris Sud-XI / INRIA

## Program Committee

George Almasi

IBM, USA

Ranieri Baraglia

CNUCE Institute, Italy

Richard Barrett

ORNL, USA

Gil Bloch

Mellanox, Israel

George Bosilca

Univeristy of Tennesse, USA

Hakon Bugge

Scali, Norway

Franck Cappello

University of Paris-Sud, France

Barbara Chapman

University of Houston, USA

Brian Coghlan

Trinity College Dublin, Ireland

Yiannis Cotronis

University of Athens, Greece

Jose Cunha

New University of Lisbon, Portugal

Marco Danelutto

University of Pisa, Italy

Luiz DeRose

Cray, USA

Frederic Desprez

INRIA, France

Erik D'Hollander

University of Ghent, Belgium

Beniamino Di Martino

Second University of Naples, Italy

Jack Dongarra

University of Tennessee, USA

Edgar Gabriel

University of Houston, USA

Al Geist

OakRidge National Laboratory, USA

Patrick Geoffray

Myricom, USA

Michael Gerndt

Technical University of Munich, Germany

Andrzej Goscinski

Deakin University, Australia

Richard L. Graham

ORNL, USA

William Gropp

Argonne National Laboratory, USA

Rolf Hempel

DLR - German Aerospace Center, Germany

## VIII Organization

|                        |                                             |
|------------------------|---------------------------------------------|
| Thomas Herault         | Université Paris Sud / INRIA, France        |
| Yutaka Ishikawa        | University of Tokyo, Japan                  |
| Rainer Keller          | HLRS, Germany                               |
| Stefan Lankes          | RWTH Aachen, Germany                        |
| Erwin Laure            | CERN, Switzerland                           |
| Laurent Lefevre        | INRIA, France                               |
| Greg Lindahl           | Pathscale, USA                              |
| Thomas Ludwig          | University of Heidelberg, Germany           |
| Ewing Rusty Lusk       | Argonne National Laboratory, USA            |
| Tomas Margalef         | Universitat Autònoma de Barcelona, Spain    |
| Jean-François Mhaut    | IMAG, France                                |
| Bernd Mohr             | Forschungszentrum Jülich, Germany           |
| Matthias Müller        | Dresden University of Technology, Germany   |
| Raymond Namyst         | University of Bordeaux, France              |
| Salvatore Orlando      | University of Venice, Italy                 |
| Christian Perez        | IRISA, France                               |
| Fabrizio Petrini       | PNNL, USA                                   |
| Neil Pundit            | Sandia National Laboratories, USA           |
| Rolf Rabenseifner      | HLRS, Germany                               |
| Thomas Rauber          | Universität Bayreuth, Germany               |
| Casiano Rodríguez-León | University of La Laguna, Spain              |
| Martin Schulz          | Lawrence Livermore National Laboratory, USA |
| Jeffrey Squyres        | Cisco, Inc., USA                            |
| Bernard Tourancheau    | University of Lyon / INRIA, France          |
| Jesper Larsson Träff   | C&C Research Labs, NEC Europe, Germany      |
| Carsten Trinitis       | Technische Universität München, Germany     |
| Roland Wismüller       | University Siegen, Germany                  |
| Felix Wolf             | Forschungszentrum Jülich, Germany           |
| Joachim Wörtingen      | C&C Research Labs, NEC Europe, Germany      |

## External Referees

|                   |                                             |
|-------------------|---------------------------------------------|
| Toni Cortes       | Universitat Politècnica de Catalunya, Spain |
| Pierre Lemarinier | Université Paris Sud/INRIA, France          |

## Conference Organization

|                  |                                 |
|------------------|---------------------------------|
| Franck Cappello  | INRIA                           |
| Gaëlle Dorkeld   | INRIA                           |
| Catherine Girard | INRIA                           |
| Chantal Girodon  | INRIA                           |
| Thomas Herault   | Université Paris Sud-XI / INRIA |

## Sponsors

The conference would have been significantly more expensive and much less pleasant to organize without the generous support of our industrial sponsors. Platinum and Gold level sponsors also gave talks at the vendor session on their latest products in parallel systems and message passing softwares. EuroPVM/MPI 2007 gratefully acknowledges the contributions of the sponsors to a successful conference.

### Platinum Level Sponsors




### Gold Level Sponsors





### Standard Level Sponsors



# Lecture Notes in Computer Science

## Sublibrary 2: Programming and Software Engineering

For information about Vols. 1– 4079  
please contact your bookseller or Springer

Vol. 4764: P. Abrahamsson, N. Baddoo, T. Margaria, R. Messnarz (Eds.), *Software Process Improvement*. XI, 225 pages. 2007.

Vol. 4758: F. Oquendo (Ed.), *Software Architecture*. XVI, 340 pages. 2007.

Vol. 4757: F. Cappello, T. Herault, J. Dongarra (Eds.), *Recent Advances in Parallel Virtual Machine and Message Passing Interface*. XVI, 396 pages. 2007.

Vol. 4753: E. Duval, R. Klamma, M. Wolpers (Eds.), *Creating New Learning Experiences on a Global Scale*. XII, 518 pages. 2007.

Vol. 4749: B.J. Krämer, K.-J. Lin, P. Narasimhan (Eds.), *Service-Oriented Computing – ICSOC 2007*. XIX, 629 pages. 2007.

Vol. 4748: K. Wolter (Ed.), *Formal Methods and Stochastic Models for Performance Evaluation*. X, 301 pages. 2007.

Vol. 4741: C. Bessière (Ed.), *Principles and Practice of Constraint Programming – CP 2007*. XV, 890 pages. 2007.

Vol. 4735: G. Engels, B. Opdyke, D.C. Schmidt, F. Weil (Eds.), *Model Driven Engineering Languages and Systems*. XV, 698 pages. 2007.

Vol. 4680: F. Saglietti, N. Oster (Eds.), *Computer Safety, Reliability, and Security*. XV, 548 pages. 2007.

Vol. 4670: V. Dahl, I. Niemelä (Eds.), *Logic Programming*. XII, 470 pages. 2007.

Vol. 4634: H. Riis Nielson, G. Filé (Eds.), *Static Analysis*. XI, 469 pages. 2007.

Vol. 4615: R. de Lemos, C. Gacek, A. Romanovsky (Eds.), *Architecting Dependable Systems IV*. XIV, 435 pages. 2007.

Vol. 4610: B. Xiao, L.T. Yang, J. Ma, C. Muller-Schloer, Y. Hua (Eds.), *Autonomic and Trusted Computing*. XVIII, 571 pages. 2007.

Vol. 4609: E. Ernst (Ed.), *ECOOP 2007 – Object-Oriented Programming*. XIII, 625 pages. 2007.

Vol. 4608: H.W. Schmidt, I. Crnković, G.T. Heineman, J.A. Stafford (Eds.), *Component-Based Software Engineering*. XII, 283 pages. 2007.

Vol. 4591: J. Davies, J. Gibbons (Eds.), *Integrated Formal Methods*. IX, 660 pages. 2007.

Vol. 4589: J. Münch, P. Abrahamsson (Eds.), *Product-Focused Software Process Improvement*. XII, 414 pages. 2007.

Vol. 4574: J. Derrick, J. Vain (Eds.), *Formal Techniques for Networked and Distributed Systems – FORTE 2007*. XI, 375 pages. 2007.

Vol. 4556: C. Stephanidis (Ed.), *Universal Access in Human-Computer Interaction*, Part III. XXII, 1020 pages. 2007.

Vol. 4555: C. Stephanidis (Ed.), *Universal Access in Human-Computer Interaction*, Part II. XXII, 1066 pages. 2007.

Vol. 4554: C. Stephanidis (Ed.), *Universal Access in Human Computer Interaction*, Part I. XXII, 1054 pages. 2007.

Vol. 4553: J.A. Jacko (Ed.), *Human-Computer Interaction*, Part IV. XXIV, 1225 pages. 2007.

Vol. 4552: J.A. Jacko (Ed.), *Human-Computer Interaction*, Part III. XXI, 1038 pages. 2007.

Vol. 4551: J.A. Jacko (Ed.), *Human-Computer Interaction*, Part II. XXIII, 1253 pages. 2007.

Vol. 4550: J.A. Jacko (Ed.), *Human-Computer Interaction*, Part I. XXIII, 1240 pages. 2007.

Vol. 4542: P. Sawyer, B. Paech, P. Heymans (Eds.), *Requirements Engineering: Foundation for Software Quality*. IX, 384 pages. 2007.

Vol. 4536: G. Concas, E. Damiani, M. Scotto, G. Succì (Eds.), *Agile Processes in Software Engineering and Extreme Programming*. XV, 276 pages. 2007.

Vol. 4530: D.H. Akehurst, R. Vogel, R.F. Paige (Eds.), *Model Driven Architecture - Foundations and Applications*. X, 219 pages. 2007.

Vol. 4523: Y.-H. Lee, H.-N. Kim, J. Kim, Y.W. Park, L.T. Yang, S.W. Kim (Eds.), *Embedded Software and Systems*. XIX, 829 pages. 2007.

Vol. 4498: N. Abdennahder, F. Kordon (Eds.), *Reliable Software Technologies - Ada-Europe 2007*. XII, 247 pages. 2007.

Vol. 4486: M. Bernardo, J. Hillston (Eds.), *Formal Methods for Performance Evaluation*. VII, 469 pages. 2007.

Vol. 4470: Q. Wang, D. Pfahl, D.M. Raffo (Eds.), *Software Process Dynamics and Agility*. XI, 346 pages. 2007.

Vol. 4468: M.M. Bonsangue, E.B. Johnsen (Eds.), *Formal Methods for Open Object-Based Distributed Systems*. X, 317 pages. 2007.

Vol. 4467: A.L. Murphy, J. Vitek (Eds.), *Coordination Models and Languages*. X, 325 pages. 2007.

Vol. 4454: Y. Gurevich, B. Meyer (Eds.), *Tests and Proofs*. IX, 217 pages. 2007.

Vol. 4444: T. Reps, M. Sagiv, J. Bauer (Eds.), *Program Analysis and Compilation, Theory and Practice*. X, 361 pages. 2007.

Vol. 4440: B. Liblit, *Cooperative Bug Isolation*. XV, 101 pages. 2007.

Vol. 4408: R. Choren, A. Garcia, H. Giese, H.-f. Leung, C. Lucena, A. Romanovsky (Eds.), *Software Engineering for Multi-Agent Systems V*. XII, 233 pages. 2007.

Vol. 4406: W. De Meuter (Ed.), *Advances in Smalltalk*. VII, 157 pages. 2007.

Vol. 4405: L. Padgham, F. Zambonelli (Eds.), *Agent-Oriented Software Engineering VII*. XII, 225 pages. 2007.

Vol. 4401: N. Guelfi, D. Buchs (Eds.), *Rapid Integration of Software Engineering Techniques*. IX, 177 pages. 2007.

Vol. 4385: K. Coninx, K. Luyten, K.A. Schneider (Eds.), *Task Models and Diagrams for Users Interface Design*. XI, 355 pages. 2007.

Vol. 4383: E. Bin, A. Ziv, S. Ur (Eds.), *Hardware and Software, Verification and Testing*. XII, 235 pages. 2007.

Vol. 4379: M. Südholt, C. Consel (Eds.), *Object-Oriented Technology*. VIII, 157 pages. 2007.

Vol. 4364: T. Kühne (Ed.), *Models in Software Engineering*. XI, 332 pages. 2007.

Vol. 4355: J. Julliand, O. Kouchnarenko (Eds.), *B 2007: Formal Specification and Development in B*. XIII, 293 pages. 2006.

Vol. 4354: M. Hanus (Ed.), *Practical Aspects of Declarative Languages*. X, 335 pages. 2006.

Vol. 4350: M. Clavel, F. Durán, S. Eker, P. Lincoln, N. Martí-Ollet, J. Meseguer, C. Talcott, *All About Maude - A High-Performance Logical Framework*. XXII, 797 pages. 2007.

Vol. 4348: S. Tucker Taft, R.A. Duff, R.L. Brukardt, E. Plödereder, P. Leroy, *Ada 2005 Reference Manual*. XXII, 765 pages. 2006.

Vol. 4346: L. Brim, B. Haverkort, M. Leucker, J. van de Pol (Eds.), *Formal Methods: Applications and Technology*. X, 363 pages. 2007.

Vol. 4344: V. Gruhn, F. Oquendo (Eds.), *Software Architecture*. X, 245 pages. 2006.

Vol. 4340: R. Prodan, T. Fahringer, *Grid Computing*. XXIII, 317 pages. 2007.

Vol. 4336: V.R. Basili, D. Rombach, K. Schneider, B. Kitchenham, D. Pfahl, R.W. Selby (Eds.), *Empirical Software Engineering Issues*. XVII, 193 pages. 2007.

Vol. 4326: S. Göbel, R. Malkewitz, I. Iurgel (Eds.), *Technologies for Interactive Digital Storytelling and Entertainment*. X, 384 pages. 2006.

Vol. 4323: G. Doherty, A. Blandford (Eds.), *Interactive Systems*. XI, 269 pages. 2007.

Vol. 4322: F. Kordon, J. Sztipanovits (Eds.), *Reliable Systems on Unreliable Networked Platforms*. XIV, 317 pages. 2007.

Vol. 4309: P. Inverardi, M. Jazayeri (Eds.), *Software Engineering Education in the Modern Age*. VIII, 207 pages. 2006.

Vol. 4294: A. Dan, W. Lamersdorf (Eds.), *Service-Oriented Computing – ICSOC 2006*. XIX, 653 pages. 2006.

Vol. 4290: M. van Steen, M. Henning (Eds.), *Middleware 2006*. XIII, 425 pages. 2006.

Vol. 4279: N. Kobayashi (Ed.), *Programming Languages and Systems*. XI, 423 pages. 2006.

Vol. 4262: K. Havelund, M. Núñez, G. Roşu, B. Wolff (Eds.), *Formal Approaches to Software Testing and Runtime Verification*. VIII, 255 pages. 2006.

Vol. 4260: Z. Liu, J. He (Eds.), *Formal Methods and Software Engineering*. XII, 778 pages. 2006.

Vol. 4257: I. Richardson, P. Runeson, R. Messnarz (Eds.), *Software Process Improvement*. XI, 219 pages. 2006.

Vol. 4242: A. Rashid, M. Aksit (Eds.), *Transactions on Aspect-Oriented Software Development II*. IX, 289 pages. 2006.

Vol. 4229: E. Najm, J.-F. Pradat-Peyre, V.V. Donzeau-Gouge (Eds.), *Formal Techniques for Networked and Distributed Systems - FORTE 2006*. X, 486 pages. 2006.

Vol. 4227: W. Nejdl, K. Tochtermann (Eds.), *Innovative Approaches for Learning and Knowledge Sharing*. XVII, 721 pages. 2006.

Vol. 4218: S. Graf, W. Zhang (Eds.), *Automated Technology for Verification and Analysis*. XIV, 540 pages. 2006.

Vol. 4214: C. Hofmeister, I. Crnković, R. Reussner (Eds.), *Quality of Software Architectures*. X, 215 pages. 2006.

Vol. 4204: F. Benhamou (Ed.), *Principles and Practice of Constraint Programming - CP 2006*. XVIII, 774 pages. 2006.

Vol. 4199: O. Nierstrasz, J. Whittle, D. Harel, G. Reggio (Eds.), *Model Driven Engineering Languages and Systems*. XVI, 798 pages. 2006.

Vol. 4192: B. Mohr, J.L. Träff, J. Worringer, J. Dongarra (Eds.), *Recent Advances in Parallel Virtual Machine and Message Passing Interface*. XVI, 414 pages. 2006.

Vol. 4184: M. Bravetti, M. Núñez, G. Zavattaro (Eds.), *Web Services and Formal Methods*. X, 289 pages. 2006.

Vol. 4166: J. Górski (Ed.), *Computer Safety, Reliability, and Security*. XIV, 440 pages. 2006.

Vol. 4158: L.T. Yang, H. Jin, J. Ma, T. Ungerer (Eds.), *Autonomic and Trusted Computing*. XIV, 613 pages. 2006.

Vol. 4157: M. Butler, C.B. Jones, A. Romanovsky, E. Troubitsyna (Eds.), *Rigorous Development of Complex Fault-Tolerant Systems*. X, 403 pages. 2006.

Vol. 4143: R. Lämmel, J. Saraiva, J. Visser (Eds.), *Generative and Transformational Techniques in Software Engineering*. X, 471 pages. 2006.

Vol. 4134: K. Yi (Ed.), *Static Analysis*. XIII, 443 pages. 2006.

Vol. 4119: C. Dony, J.L. Knudsen, A. Romanovsky, A.R. Tripathi (Eds.), *Advanced Topics in Exception Handling Techniques*. X, 302 pages. 2006.

Vol. 4111: F.S. de Boer, M.M. Bonsangue, S. Graf, W.-P. de Roever (Eds.), *Formal Methods for Components and Objects*. VIII, 447 pages. 2006.

Vol. 4089: W. Löwe, M. Südholt (Eds.), *Software Composition*. X, 339 pages. 2006.

Vol. 4085: J. Misra, T. Nipkow, E. Sekerinski (Eds.), *FM 2006: Formal Methods*. XV, 620 pages. 2006.

# Table of Contents

## Invited Talks

|                                                                                                                  |    |
|------------------------------------------------------------------------------------------------------------------|----|
| The X-Scale Challenge.....                                                                                       | 1  |
| <i>George Bosilca</i>                                                                                            |    |
| Sustained Petascale: The Next MPI Challenge .....                                                                | 3  |
| <i>Al Geist</i>                                                                                                  |    |
| MPI: Past, Present and Future .....                                                                              | 5  |
| <i>Tony Hey</i>                                                                                                  |    |
| New and Old Tools and Programming Models for High-Performance<br>Computing.....                                  | 7  |
| <i>Ewing Lusk</i>                                                                                                |    |
| The TSUBAME Cluster Experience a Year Later, and Onto Petascale<br>TSUBAME 2.0 .....                             | 8  |
| <i>Satoshi Matsuoka</i>                                                                                          |    |
| To Infinity and Beyond?! On Scaling Performance Measurement and<br>Analysis Tools for Parallel Programming ..... | 10 |
| <i>Bernd Mohr</i>                                                                                                |    |

## Tutorials

|                                                |    |
|------------------------------------------------|----|
| Using MPI-2: A Problem-Based Approach.....     | 12 |
| <i>William D. Gropp and Ewing Lusk</i>         |    |
| Verifying Parallel Programs with MPI-Spin..... | 13 |
| <i>Stephen F. Siegel</i>                       |    |
| Advanced MPI Programming .....                 | 15 |
| <i>Julien Langou and George Bosilca</i>        |    |

## Outstanding Papers

|                                                                           |    |
|---------------------------------------------------------------------------|----|
| Full Bandwidth Broadcast, Reduction and Scan with Only Two<br>Trees ..... | 17 |
| <i>Peter Sanders, Jochen Speck, and Jesper Larsson Träff</i>              |    |
| Process Cooperation in Multiple Message Broadcast .....                   | 27 |
| <i>Bin Jia</i>                                                            |    |

|                                                                                                                 |     |
|-----------------------------------------------------------------------------------------------------------------|-----|
| Self-consistent MPI Performance Requirements .....                                                              | 36  |
| <i>Jesper Larsson Träff, William Gropp, and Rajeev Thakur</i>                                                   |     |
| Test Suite for Evaluating Performance of MPI Implementations That Support MPLTHREAD_MULTIPLE .....              | 46  |
| <i>Rajeev Thakur and William Gropp</i>                                                                          |     |
| <b>Applications</b>                                                                                             |     |
| An Improved Parallel XSL-FO Rendering for Personalized Documents .....                                          | 56  |
| <i>Luiz Gustavo Fernandes, Thiago Nunes, Mateus Raeder, Fabio Giannetti, Alexis Cabeda, and Guilherme Bedin</i> |     |
| An Extensible Framework for Distributed Testing of MPI Implementations .....                                    | 64  |
| <i>Joshua Hursey, Ethan Mallove, Jeffrey M. Squires, and Andrew Lumsdaine</i>                                   |     |
| A Virtual Test Environment for MPI Development: Quick Answers to Many Small Questions .....                     | 73  |
| <i>Wolfgang Schnerring, Christian Kauhaus, and Dietmar Fey</i>                                                  |     |
| Multithreaded Tomographic Reconstruction .....                                                                  | 81  |
| <i>José Antonio Álvarez, Javier Roca, and Jose Jesús Fernández</i>                                              |     |
| Parallelizing Dense Linear Algebra Operations with Task Queues in 11c .....                                     | 89  |
| <i>Antonio J. Dorta, José M. Badía, Enrique S. Quintana-Ortí, and Francisco de Sande</i>                        |     |
| ParaLEX: A Parallel Extension for the CPLEX Mixed Integer Optimizer .....                                       | 97  |
| <i>Yuji Shinano and Tetsuya Fujie</i>                                                                           |     |
| Performance Analysis and Tuning of the XNS CFD Solver on Blue Gene/L .....                                      | 107 |
| <i>Brian J.N. Wylie, Markus Geimer, Mike Nicolai, and Markus Probst</i>                                         |     |
| (Sync Async) <sup>+</sup> MPI Search Engines .....                                                              | 117 |
| <i>Mauricio Marin and Veronica Gil Costa</i>                                                                    |     |

## Collective Operations

|                                                                                                     |     |
|-----------------------------------------------------------------------------------------------------|-----|
| A Case for Standard Non-blocking Collective Operations .....                                        | 125 |
| <i>Torsten Hoefler, Prabhanjan Kambadur, Richard L. Graham, Galen Shipman, and Andrew Lumsdaine</i> |     |

|                                                                                                      |     |
|------------------------------------------------------------------------------------------------------|-----|
| Optimization of Collective Communications in HeteroMPI .....                                         | 135 |
| <i>Alexey Lastovetsky, Maureen O'Flynn, and Vladimir Rychkov</i>                                     |     |
| <b>Fault Tolerance</b>                                                                               |     |
| Low Cost Self-healing in MPI Applications .....                                                      | 144 |
| <i>Jacques A. da Silva and Vinod E.F. Rebello</i>                                                    |     |
| Fault Tolerant File Models for MPI-IO Parallel File Systems .....                                    | 153 |
| <i>A. Calderón, F. García-Carballeira, Florin Isailă, Rainer Keller, and Alexander Schulz</i>        |     |
| <b>Library Internals</b>                                                                             |     |
| An Evaluation of Open MPI's Matching Transport Layer on the Cray XT .....                            | 161 |
| <i>Richard L. Graham, Ron Brightwell, Brian Barrett, George Bosilca, and Jelena Pješivac-Grbović</i> |     |
| Improving Reactivity and Communication Overlap in MPI Using a Generic I/O Manager .....              | 170 |
| <i>François Trahay, Alexandre Denis, Olivier Aumage, and Raymond Namyst</i>                          |     |
| Investigations on InfiniBand: Efficient Network Buffer Utilization at Scale .....                    | 178 |
| <i>Galen M. Shipman, Ron Brightwell, Brian Barrett, Jeffrey M. Squyres, and Gil Bloch</i>            |     |
| <b>Message Passing on Hierarchical Machines and Grids</b>                                            |     |
| Improving MPI Support for Applications on Hierarchically Distributed Resources .....                 | 187 |
| <i>Raúl López and Christian Pérez</i>                                                                |     |
| MetaLoRaS: A Re-scheduling and Prediction MetaScheduler for Non-dedicated Multiclusters .....        | 195 |
| <i>J.Ll. Lérída, F. Solsona, F. Giné, M. Hanzich, J.R. García, and P. Hernández</i>                  |     |
| Using CMT in SCTP-Based MPI to Exploit Multiple Interfaces in Cluster Nodes .....                    | 204 |
| <i>Brad Penoff, Mike Tsai, Janardhan Iyengar, and Alan Wagner</i>                                    |     |
| <b>MPI-I/O</b>                                                                                       |     |
| Analysis of the MPI-IO Optimization Levels with the PIOViz Jumpshot Enhancement .....                | 213 |
| <i>Thomas Ludwig, Stephan Krempel, Michael Kuhn, Julian Kunkel, and Christian Lohse</i>              |     |

|                                                                     |     |
|---------------------------------------------------------------------|-----|
| Extending the MPI-2 Generalized Request Interface .....             | 223 |
| <i>Robert Latham, William Gropp, Robert Ross, and Rajeev Thakur</i> |     |
| Transparent Log-Based Data Storage in MPI-IO Applications .....     | 233 |
| <i>Dries Kimpe, Rob Ross, Stefan Vandewalle, and Stefaan Poedts</i> |     |

**One-Sided**

|                                                                                                         |     |
|---------------------------------------------------------------------------------------------------------|-----|
| Analysis of Implementation Options for MPI-2 One-Sided .....                                            | 242 |
| <i>Brian W. Barrett, Galen M. Shipman, and Andrew Lumsdaine</i>                                         |     |
| MPI-2 One-Sided Usage and Implementation for Read Modify Write Operations: A Case Study with HPCC ..... | 251 |
| <i>Gopalakrishnan Santhanaraman, Sundeep Narravula, Amith.R. Mamidala, and Dhabaleswar K. Panda</i>     |     |
| RDMA in the SiCortex Cluster Systems .....                                                              | 260 |
| <i>Lawrence C. Stewart, David Gingold, Jud Leonard, and Peter Watkins</i>                               |     |
| Revealing the Performance of MPI RMA Implementations .....                                              | 272 |
| <i>William D. Gropp and Rajeev Thakur</i>                                                               |     |

**PVM and Harness**

|                                                                                                        |     |
|--------------------------------------------------------------------------------------------------------|-----|
| Distributed Real-Time Computing with Harness .....                                                     | 281 |
| <i>Emanuele Di Saverio, Marco Cesati, Christian Di Biagio, Guido Pennella, and Christian Engelmann</i> |     |
| Frequent Itemset Mining with Trie Data Structure and Parallel Execution with PVM .....                 | 289 |
| <i>Levent Guner and Pinar Senkul</i>                                                                   |     |

**Tools**

|                                                                                                 |     |
|-------------------------------------------------------------------------------------------------|-----|
| Retrospect: Deterministic Replay of MPI Applications for Interactive Distributed Debugging..... | 297 |
| <i>Aurelien Bouteiller, George Bosilca, and Jack Dongarra</i>                                   |     |
| Extended MPICC to Generate MPI Derived Datatypes from C Datatypes Automatically .....           | 307 |
| <i>Éric Renault</i>                                                                             |     |
| Timestamp Synchronization for Event Traces of Large-Scale Message-Passing Applications .....    | 315 |
| <i>Daniel Becker, Rolf Rabenseifner, and Felix Wolf</i>                                         |     |

## Verification of Message Passing Programs

|                                                                                                               |     |
|---------------------------------------------------------------------------------------------------------------|-----|
| Verification of Halting Properties for MPI Programs Using Nonblocking Operations .....                        | 326 |
| <i>Stephen F. Siegel and George S. Avrunin</i>                                                                |     |
| Correctness Debugging of Message Passing Programs Using Model Verification Techniques .....                   | 335 |
| <i>Robert Lovas and Peter Kacsuk</i>                                                                          |     |
| Practical Model-Checking Method for Verifying Correctness of MPI Programs .....                               | 344 |
| <i>Salman Pervez, Ganesh Gopalakrishnan, Robert M. Kirby, Robert Palmer, Rajeev Thakur, and William Gropp</i> |     |

## ParSim

|                                                                                                                                                                          |     |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 6 <sup>th</sup> International Special Session on Current Trends in Numerical Simulation for Parallel Engineering Environments: New Directions and Work-in-Progress ..... | 354 |
| <i>Martin Schulz and Carsten Trinitis</i>                                                                                                                                |     |
| Gyrokinetic Semi-lagrangian Parallel Simulation Using a Hybrid OpenMP/MPI Programming .....                                                                              | 356 |
| <i>G. Latu, N. Crouseilles, V. Grandgirard, and E. Sonnendrücker</i>                                                                                                     |     |
| Automatic Parallelization of Object Oriented Models Executed with Inline Solvers .....                                                                                   | 365 |
| <i>Håkan Lundvall and Peter Fritzson</i>                                                                                                                                 |     |
| 3D Parallel Elastodynamic Modeling of Large Subduction Earthquakes .....                                                                                                 | 373 |
| <i>Eduardo Cabrera, Mario Chavez, Raúl Madariaga, Narciso Perea, and Marco Frisenda</i>                                                                                  |     |

## Posters Abstracts

|                                                                                             |     |
|---------------------------------------------------------------------------------------------|-----|
| Virtual Parallel Machines Through Virtualization: Impact on MPI Executions .....            | 381 |
| <i>Benjamin Quetier, Thomas Herault, Vincent Neri, and Franck Cappello</i>                  |     |
| Seshat Collects MPI Traces: Extended Abstract .....                                         | 384 |
| <i>Rolf Riesen</i>                                                                          |     |
| Dynamic Optimization of Load Balance in MPI Broadcast .....                                 | 387 |
| <i>Takesi Soga, Kouji Kurihara, Takeshi Nanri, Motoyoshi Kurokawa, and Kazuaki Murakami</i> |     |

|                                                                                                           |     |
|-----------------------------------------------------------------------------------------------------------|-----|
| An Empirical Study of Optimization in Seamless Remote MPI-I/O for<br>Long Latency Network .....           | 389 |
| <i>Yuichi Tsujita</i>                                                                                     |     |
| Multithreaded and Distributed Simulation of Large Biological Neuronal<br>Networks .....                   | 391 |
| <i>Jochen M. Eppler, Hans E. Plesser, Abigail Morrison,<br/>Markus Diesmann, and Marc-Oliver Gewaltig</i> |     |
| Grid Services for MPI .....                                                                               | 393 |
| <i>Camille Coti, Ala Rezmerita, Thomas Herault, and Franck Cappello</i>                                   |     |
| <b>Author Index</b> .....                                                                                 | 395 |