
IWCC'99

blank

Proceedings

1st IEEE Computer Society International Workshop on Cluster Computing

2-3 December 1999

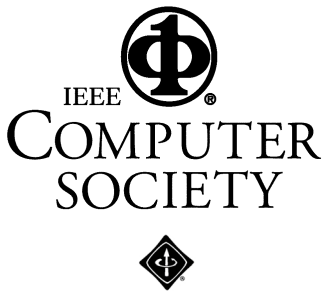
Melbourne, Australia

Sponsored by

IEEE Computer Society
IEEE Computer Society Task Force on Cluster Computing
Monash University

Edited by

Rajkumar Buyya
Mark Baker
Ken Hawick
Heath James



Los Alamitos, California
Washington · Brussels · Tokyo

IEEE Computer Society Order Number PR00343
ISBN 0-7695-0343-8
ISBN 0-7695-0345-4 (microfiche)
Library of Congress Number 99-64613

Additional copies may be ordered from:

IEEE Computer Society
Customer Service Center
10662 Los Vaqueros Circle
P.O. Box 3014
Los Alamitos, CA 90720-1314
Tel: + 1-714-821-8380
Fax: + 1-714-821-4641
E-mail: cs.books@computer.org

IEEE Service Center
445 Hoes Lane
P.O. Box 1331
Piscataway, NJ 08855-1331
Tel: + 1-732-981-0060
Fax: + 1-732-981-9667
mis.custserv@computer.org

IEEE Computer Society
Asia/Pacific Office
Watanabe Building,
1-4-2 Minami-Aoyama
Minato-ku, Tokyo 107-0062 JAPAN
Tel: + 81-3-3408-3118
Fax: + 81-3-3408-3553
tokyo.ofc@computer.org

Editorial production by Danielle C. Martin

Cover art production by Joseph Daigle/Studio Productions

Printed in the United States of America by The Printing House

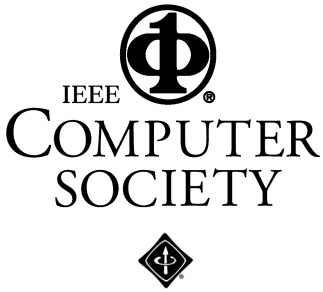


Table of Contents

1st IEEE Computer Society International Workshop on Cluster Computing (IWCC'99)

CLUSTER PROGRAMMING AND ANALYSIS MODELS

ALGORITHMS AND APPLICATIONS

INDUSTRY TRACK

Foreword

It is the best of computing, it is the worst of computing. But in both cases, Cluster Computing in general and PC Clusters in particular have become a dominant force in the rapid evolution of mid-range and high end computing. It is therefore both an honor and a pleasure for me to introduce this important new forum, the 1st IEEE International Workshop on Clustered Computing, which is likely to have a strong influence on the growth and impact of the emerging interdisciplinary field of cluster computing. Permit me, then, to take this opportunity to share with you my views on the extraordinary opportunities and the crucial unresolved issues that characterize the domain of clusters implementation and application, derived from my experiences while engaged in research related to Beowulf-class systems.

Linux based Beowulf-class PC clusters, NT-PC clusters, clusters of workstations, and even clusters of commercial DSMs have evolved as the principal path to scalability beyond the mainstream commercial marketing sweet spot of desktops, file servers, and mid-range compute servers. But it is the low cost Beowulf-class systems and other families of PC clusters that have emerged as the driver of innovation in supercomputing by breaking the price-performance barrier, delivering order of magnitude advantage to appropriate applications, such that "it is the best of computing" as recognized by the 1997 and 1998 Gorden Bell Prize for supercomputing price-performance. Beyond cost, ensembles of mass market processor and networking devices provide great freedom of configuration, enhancement, and software diversity. Benefiting from the enormous accrued talent and effort being applied to Linux and its base of cooperative parallel processing system software and tools with open source licensing, Beowulf-class systems and their siblings have become a common platform derived from widely available hardware and software building blocks.

The resulting level of confidence by the applications community in the likely stability and longevity of this class of parallel architecture and programming model is unprecedented in the history of supercomputing. No longer can the marketing decisions of a single vendor or software house completely disrupt the productivity of the applications programmers or users of their software products. Similarly, clusters offer a single generic model to which system software research groups and commercial enterprises can target their efforts with the confidence of a broad and increasing consumer base.

But far from being a direct competitor to commercial vendors, Beowulf provides a new business opportunity for vendors to both broaden their product lines and significantly increase their customer base by offering a low entry-level cost family of parallel systems. An important side effect of this is the significant expansion of the number of parallel programmers and the scope of parallel applications, which will increase the acceptance of other vendor parallel computer products. In the last year, a number of the major computer manufacturers including SGI, IBM, Compaq, and others have made important announcements about their own Beowulf offerings and their future use of the Linux operating system. Large PC clusters including those incorporating the advanced Compaq Alpha microprocessor have entered the "Top 500" list of the world's most powerful computers. It is likely that an increasing percentage of the entries on this list will be Beowulf-class systems of ever increasing number of nodes. Also in the last year, two important books have been published highlighting the nature and means of accomplishing real-world computation with clusters of PCs.

And yet, important limitations of Beowulfs and other classes of clusters have been exposed with the rapid growth of the installed base and use of clusters as well as their increase in average number of nodes, range of application workloads, and user environments. When Donald Becker and I undertook the first Beowulf Project in 1994, clusters of at most a few dozen single-processor nodes were applied to one or a few problems in a small localized environment of a small number of users. The hardware and software methodologies we employed were sufficient and many of the successes with Beowulfs were achieved through these means. But this "mom-and-pop" approach to do-it-yourself supercomputing is no longer a viable path to effective community expansion. These early software tools are proving inadequate to support the more sophisticated user environments such that in some users eyes "it is the worst of computing". Spartan software combined with anemic networks exhibiting the combined properties of lower bandwidth and longer latencies are obstacles to wider usage of this otherwise extraordinarily cost effective approach to high end computing. Currently in research labs and commercial sites, advances in both hardware and software are addressing many of these key challenges in conjunction with the development of novel approaches to latency tolerant algorithm design. These innovations may promise a new generation of sophisticated, robust, and flexible clustered computers that will relegate the early Beowulfs to the pages of pioneering history.

It is at this critical period in the rapid evolution of PC cluster computing technologies that the first International Workshop of Cluster Computing is being inaugurated. Under the sponsorship of the IEEE and its Task Force on Cluster Computing, this new conference is a valuable forum for the sharing of results and the exchange of ideas that will provide critical impetus to future directions in research and development. As an international meeting, it embraces the world community and contributions being made globally, reflecting that the creation and use of Beowulfs and other PC clusters is truly a planet-wide phenomenon comprising the concepts and talents that transcend conventional boundaries. The proceedings of this conference includes the ideas and findings of many current contributors from many nations cooperating as a distributed collegial confederation bound by a common vision, motivated by shared computing requirements and enabled through world-wide communications, commerce, and technical opportunity. It is with great pleasure that I congratulate the organizers for establishing this new series of conferences and that I welcome you to a next step in one of today's most exciting and important fields of computing.

Thomas Sterling

California Institute of Technology and
NASA Jet Propulsion Laboratory, California, USA

Message from the General Co-Chairs

Welcome to the first IEEE Task Force on Cluster Computing (TFCC) sponsored event. The International Workshop on Cluster Computing (IWCC'99) is sponsored by the IEEE Computer Society through the TFCC with generous contributions from Hewlett Packard (USA), MPI Software Technology Inc. (USA), Asian Technology Information Program (ATIP, Japan), GENIAS Software GmbH (Germany), Sun Microsystems Inc. (USA), Compaq Computer Australia, Distributed Systems Technology Centre (Australia), and Monash University (Australia). We wish to express our sincere gratitude to all our supporters.

The Task Force itself was approved in late December 1998 and commenced its activities in February 1999. The TFCC sponsors professional meetings, brings out publications, sets guidelines for educational programs, and helps to co-ordinate academic, funding agency, and industrial activities. Since its inception the TFCC has been actively working on initiating its program for the benefit of cluster computing community. These activities are undertaken with the support of a group of volunteer members who comprise our executive and advisory committees. The TFCC has been successful in promoting cluster computing through its educational program. One aspect of this program that has been especially popular is our international book donation scheme. Here we requested that authors of cluster-related books ask their publishers to donate books to our program. In fact, the program has become so popular that the IEEE Computer Magazine editor invited us to write an article on this program — it appeared in the July 1999 issue of Computer under the Technical Activities Forum section. For further information about TFCC membership (which is free and we are always looking for new volunteers), its planned and future activities, please browse:

AU site: http://www.dgs.monash.edu.au/~rajkumar/tfcc/

UK site: http://www.dcs.port.ac.uk/~mab/tfcc/

The response to the workshop's call for papers has been excellent and we expect that attendance at the actual workshop will be equally impressive. The success of the workshop is wholly due to the hard work of the local organizing committee, publicity coordinators, and program committee. In particular Kenneth Hawick, and his associate Heath James, have put in an exceptional effort in to organizing the peer review process and selecting the highly recommended papers. We appreciate their hard work and commitment in formulating such an outstanding technical program. In addition to peer reviewed accepted papers, we also have industry track papers as part of our program. We thank our international program committee members for donating their precious time for reviewing and offering their expert comments on the papers.

The workshop was originally planned as one-day event. It was necessary to extend the workshop to two days due to the outstanding response that we received. The venue, Centra Melbourne, on the banks of the Yarra River, is located in the center of the city and near many major Melbourne attractions. We would like to thank the local organizing committee for arranging such an ideal venue. We acknowledge the support of Mahbub Hassan, Thomas C. Peachey, Jonathan Giddy, Jahan Hassan, and Shiranthi Ponniah in organizing the workshop. In addition we would like to thank the organizing committee of PART'99 for their support in hosting the workshop in conjunction with their conference.

We thank our keynote speakers Kai Hwang, Anthony Skjellum and David Abramson; and industry/invited speakers Patrick Estep, Ira Pramanick, Bruce Foster, and Matthew Bailes for sharing their expertise. We express our gratitude to Paul Roe for serving as the moderator for a panel on "Cluster Computing R&D in Australia". We would like to thank our plenary speaker Thomas Sterling for writing a foreword to the workshop proceedings.

Anthony Skjellum (MPI-SoftTech), David K. Kahaner (ATIP), Wolfgang Gentzsch (Genias), Ira Pramanick (Sun), Steve Tolnai (Compaq), David Barbagallo (DSTC), Daniel Charlu & Brian Cox (HP), and David Abramson & John Rosenberg (Monash) have received our request for financial support enthusiastically and were instrumental in getting generous donations from their respective organizations. We sincerely thank all of them.

We would like to thank David Abramson for his kind permission to use the services of Monash University finance management in handling the registration fee and supporting the involvement of the School of Computer Science and Software Engineering staff members in the organization of the workshop.

The staff members at the IEEE Computer Society have helped us cheerfully in numerous ways. Our special thanks to Tracy Woods (Volunteer Services Coordinator) who served as our first point of contact right from the submission of TFCC formation proposal to the creation of this conference series. It would have been impossible to bring out these proceedings without the timely support of IEEE CS press crew: Deborah Plummer (Manager Production), Thomas Baldwin (Proceedings Supervisor), and Danielle Martin (Production Editor). The event sponsorship was approved by Willis King (VP for Conferences and Tutorials) with the assistance from Anne Kelly (Director of Volunteer Services) and Maggie Johnson (Conference Finance Coordinator). Thank you all for your cooperation and timely assistance.

We hope that your participation in the event will create a new network of colleagues, friendship, and provide a great opportunity to see the latest developments in cluster computing research in both industry and academia.

By the way, our next event “International Conference on Cluster Computing” (Cluster’2000), a merger of IWCC and other major international workshops, is being scheduled to be held in Chemnitz (the former East Germany). We hope to see you all again in Germany next fall.

Enjoy your visit to the land of Kangaroos!

Rajkumar Buyya

School of Computer Science and Software Engineering
Monash University, Australia
http://www.dgs.monash.edu.au/~rajkumar/

Mark Baker

School of Computer Science
University of Portsmouth, UK
http://www.dcs.port.ac.uk/~mab/

Message from the Program Chair

It gives me great pleasure to present the program and this proceedings for the 1st IEEE Computer Society International Workshop on Cluster Computing (IWCC'99) which will be held on 2nd and 3rd December 1999 in Melbourne, Australia.

It is a sign of the current high levels of interest and activity in cluster computing that we have contributions from: the USA; Japan; France; Hong Kong; Taiwan; Brazil; Germany; China; Singapore; the UK; and Mexico as well as from researchers throughout Australia.

Cluster computing can perhaps be described as a fusion of the fields of parallel, high-performance and distributed computing. The history of computer science is a tale of changing fashions and trends as techniques and approaches are re-invented for new problems and become more or less viable as the technology and the economics of different computer components changes. It is likely that many of the techniques for effective programming in these other three disciplines will be re-deployed, albeit often with a new spin on them, as cluster computing is now so economically attractive. There are many exciting areas of development in cluster computing with new ideas as well as hybrids of old ones being deployed for production as well as research systems. Due mainly to the diverse range of technologies that clusters can use, it was a difficult task to arrive at a coherent program for this workshop. We have tried to group the contributed papers into nine distinct categories, although inevitably there is some overlap.

- Cluster Setup and Performance Measurement
- Cluster Communications Software and Protocols
- Network Communication Optimizations
- Cluster File System and Parallel I/O
- Scheduling Programs on Clusters
- Cluster Management and Metacomputing
- Cluster Operating Systems and Monitoring
- Cluster Programming and Analysis Models
- Algorithms and Applications

There seem to be many research projects active in all these areas worldwide, and we are fortunate to be able to present a snapshot of some of the most recent work at this workshop.

This workshop would not have taken place without the key efforts of Mark Baker and Rajkumar Buyya who have been the main driving forces behind the IEEE Task Force on Cluster Computing. It is a pleasure to acknowledge their efforts and also those of all the members of the task force, particularly those who agreed to serve as program committee members for this workshop. I also wish to extend great thanks to my former graduate student Heath James for his sterling efforts in helping organize the review process, and of course to everyone who so kindly helped review papers at short notice. Through their efforts each submitted paper to IWCC'99 was refereed by at least two independent researchers. Some submissions had as many as four referees. We were faced by some difficult choices and due to time restrictions during the conference were only able to accept around 60% of the submitted papers.

Cluster computing is a growing area of activity with many research avenues now possible for researchers who would previously not have had access to high-end computing resources. I hope these proceedings reflect this, and thanks again to all the authors for their contributions.

Ken Hawick

Program Committee Chair
The University of Adelaide
Adelaide, Australia, September 1999

IWCC'99 Crew Members

General Chairs

Mark Baker, *Portsmouth University, UK*
Rajkumar Buyya, *Monash University, Australia*

Program Chair

Ken Hawick, *Adelaide University, Australia*

Program Committee

David Abramson, *Monash University, Australia*
Hamid Arabnia, *University of Georgia, USA*
David Bader, *University of New Mexico, USA*
Mark Baker, *Portsmouth University, UK*
Ricardo Bianchini, *Federal University of Rio de Janeiro, Brazil*
Suchendra Bhandarkar, *University of Georgia, USA*
Luc Bouge, *LIP, ENS Lyon, France*
Marian Bubak, *Institute of Computer Science, Poland*
Rajkumar Buyya, *Monash University, Australia*
Giovanni Chiola, *University of Genoa, Italy*
Paul Coddington, *University of Adelaide, Australia*
Toni Cortes, *Universitat Politecnica de Catalunya, Spain*
Dave DeRoure, *University of Southampton, UK*
Joao Gabriel Silva, *Coimbra University, Portugal*
Al Geist, *Oakridge National Laboratory, USA*
Andrzej Goscinski, *Deakin University, Australia*
Wolfgang Gentzsch, *Genias GmbH, Germany*
Bill Gropp, *Argonne National Laboratory, USA*
Salim Hariri, *Arizona University, USA*
Dan Hyde, *Bucknell University, USA*
Yutaka Ishikawa, *Real World Computing Partnership, Japan*
Heath James, *University of Adelaide, Australia*
Hai Jin, *University of Hong Kong, China*
Daniel S. Katz, *Jet Propulsion Laboratory, California Institute of Technology, USA*
Chung-Ta King, *National Tsinghua University, Taiwan*
Kevin Maciunas, *University of Adelaide, Australia*
Piyush Maheshwari, *University of New South Wales, Australia*
Chris McDonald, *University of Western Australia, Australia*
John Morris, *University of Western Australia, Australia*
Marcin Paprzycki, *University of Southern Mississippi, USA*
Robert Pennington, *NCSA, USA*
Ira Pramanick, *Sun Microsystems, USA*
Radharamanan Radhakrishnan, *University of Cincinnati, USA*
Rajeev Raje, *Purdue University, USA*

Mohan Ram, *Centre for Development of Advanced Computing, India*
Wolfgang Rehm, *TU Chemnitz, Germany*
Paul Roe, *Queensland University of Technology, Australia*
Harjinder Sandhu, *York University, Canada*
Danial Saverese, *California Institute of Technology, USA*
Hong Shen, *Griffith University, Australia*
R. K. Shyamasundar, *Tata Institute of Fundamental Research, India*
Tony Skjellum, *MPI Software Technology, USA*
Thomas Sterling, *California Institute of Technology, USA*
Peter Strazdins, *Australian National University, Australia*
Chengzheng Sun, *Griffith University, Australia*
Yong-Meng Teo, *National University of Singapore, Singapore*
Putchong Uthayopas, *Kasetsart University, Thailand*
David W. Walker, *Cardiff University, UK*
Barry Wilkinson, *University of North Carolina, USA*
Albert Zomaya, *University of Western Australia, Australia*

Publications/Proceedings Chair

Rajkumar Buyya, *Monash University, Australia*

Poster Papers Chairs

Heath James, *University of Adelaide, Australia*
Hai Jin, *University of Southern California, USA*

Acknowledgments

Sponsors

IEEE Computer Society Task Force on Cluster Computing, TFCC

Monash University, Australia

Premier Contributor

Hewlett Packard, USA

Contributors

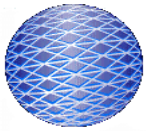
MPI Software Technology Inc., USA

Asian Technology Information Program, ATIP, Japan

GENIAS Software GmbH, Germany

Sun Microsystems Inc., USA

Compaq Computer, Australia



Distributed Systems Technology Centre, Australia