Proceedings

2001 IEEE International Conference on

Cluster Computing

Editors: Daniel S. Katz, Thomas Sterling, Mark Baker, Larry Bergman, Marcin Paprzycki, and Rajkumar Buyya

8–11 October 2001
California, USA

IEEE COMPUTER SOCIETY

IEEE

http://www.ieeetfcc.org

Sponsored by
IEEE
IEEE Computer Society and
IEEE Computer Society Task Force on Cluster Computing

TFCC
IEEE Task Force on Cluster Computing

Proceedings

2001 IEEE International Conference on Cluster Computing

8–11 October 2001

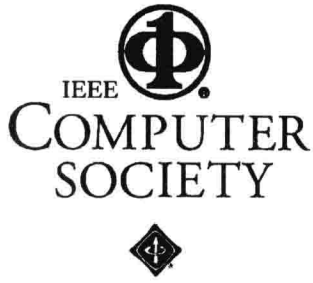
Newport Beach, California, USA

Edited by

Daniel S. Katz, Thomas Sterling, Mark Baker, Larry Bergman, Marcin Paprzycki, and Rajkumar Buyya

Sponsored by the

IEEE Computer Society Task Force on Cluster Computing



Los Alamitos, California

Washington · Brussels · Tokyo

IEEE Computer Society Order Number PR01116
ISBN 0-7695-1116-3
ISBN 0-7695-1118-X (microfiche)
Library of Congress Number 2001093692

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
http://computer.org/
csbooks@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
http://shop.ieee.org/store/
customer-service@ieee.org

IEEE Computer Society
Asia/Pacific Office
Watanabe Bldg., 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 Anne Jacobs

Cover art production by Joe Daigle/Studio Productions

Printed in the United States of America by The Printing House

Proceedings

2001 IEEE International Conference on Cluster Computing

Message from the Task Force Co-Chairs

Welcome to the third annual IEEE international cluster computing technical conference sponsored by the IEEE Computer Society's Task Force on Cluster Computing (TFCC). The conference is being held in cooperation with the Association for Computing Machinery (ACM) and hosted by the California Institute of Technology, NASA Jet Propulsion Laboratory, and University of California at Irvine.

We would like to take this opportunity to say few words about TFCC activities and acknowledge the outstanding efforts of our members and volunteers. The TFCC has been in existence since early 1999 and within its short life it has started to have an impact on the cluster computing community in academia, industry, and business. The TFCC is committed to the development of cluster computing research, development and education, both in academia and industry. Towards this end, we have undertaken a number of activities including promoting open community discussion, holding events that bring together leading experts, the instigation of a book donation programme and the provision of guidance and advise to commercial bodies. One particularly successful activity has been the TFCCs educational programme. Here we are attempting to promote the take-up of cluster-related technologies in the core curriculum of educational institutions around the world. In conjunction with this effort is our book donation programme, which has enabled the TFCC to donate nearly a thousand books to academic institutions around the world.

Earlier this year, the TFCC actively promoted the formation of a new Computing Research Repository (CoRR) dedicated to Cluster Computing literature. The aim of CoRR e-Print Archive is to promote science and technology by providing an information repository that is internationally accessible. Together with the TFCC, this activity is sponsored and promoted by the ACM, the Los Alamos e-Print archive, and Networked Computer Science Technical Reference Library.

In association with the TOP500 list organizers, the TFCC has instigated a similar new list dedicated to clusters. The new list (http://www.topclusters.org) is not just about reporting peak-performance; it also includes various other parameters including database, web, communication, and application level performance. We thank the TFCC members for contributing their time and energy to this effort during the initial discussions about the formation of this list.

The TFCCs activities are undertaken with the support of a group of active volunteer who are part of its executive and advisory committees. The executive committee comprises chairs, vice chairs, technical area coordinators, and regional area coordinators. The TFCC has a presence on all continents with backing from regional coordinators in: Africa, Australia, Canada, China, Europe, Hong Kong, India, Japan, the middle-east, South America, Thailand, Taiwan, and USA. We appreciate and acknowledge their support.

The TFCC has organised and sponsored a number of other events in addition this annual conference. Further information about membership of the TFCC, its planned and future activities can be found on our Web pages[1]. It should be noted that TFCC membership is free and we are always looking for willing volunteers to actively engage in organization and management of TFCC activities.

---

[1] TFCC – http://www.ieeetfcc.org

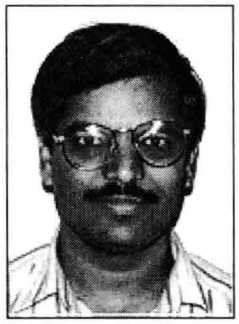
The response to our previous annual conferences[2] held in Melbourne, Australia in 1999 and Chemnitz, Germany in 2000 has been first-rate and attracted hundreds of participants from all over the world. We expect a similar response for this years' conference. The success of the conference is wholly due to the hard work of the local organising committee, the publicity co-coordinators, as well as the program and steering committees. We would like to thank Thomas Sterling, Daniel Katz, and Larry Bergman for volunteering to lead organizational efforts. Thomas and Dan managed submissions, reviews, and selection of papers with the help of program committee members. We thank members of the international programme committee for volunteering their precious time and effort.

We acknowledge Kai Hwang for his initial efforts on the organization of the conference. We appreciate the efforts of Nalini Venkatasubramanian for handling many aspects of local organizational matters. We also thank Rick Stevens for his support and advise. Our thanks also go to Ira Pramanick, Marcin Paprzycki, Hai Jin, and Tim Pinkston for serving as tutorial, proceedings, publicity, and finance chairs respectively.

We also would like to thank the IEEE Computer Society for their continued support of the TFCCs activities and for formally sponsoring this conference. We thank Tom Baldwin, the IEEE CS Proceedings Manager for helping us out with proceedings' publication contractual matters. Our special thanks also go to Anne Jacobs for managing the final proceeding publication.

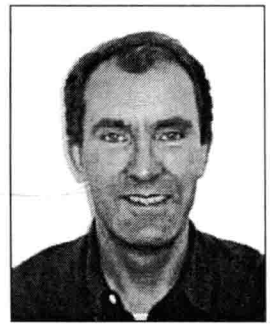
We hope that your participation in this event will help you create a new network of colleagues, engender friendship, and provide a great opportunity to see the latest developments in cluster computing in both industry and academia.

Enjoy your visit to Newport Beach, California and the USA!



Rajkumar Buyya

Monash University, Australia
http://www.buyya.com



Mark Baker

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

[2] TFCC Annual Conference - http://www.clustercomp.org

Message from the Conference General Co-Chairs

We extend a warm welcome to you all to Cluster 2001, the annual international cluster computing conference sponsored by the IEEE Computer Society's Task Force on Cluster Computing (TFCC). The event this year is being held at the Sutton Place Hotel in the beautiful location of Newport Beach, California.

The organization of a conference, such as Cluster 2001, is an onerous task and could not be accomplished without the help of a dedicated group of enthusiastic and willing volunteers.

We would like to first express a special thanks to Thomas Sterling and Dan Katz for their efforts organizing and putting together the technical content of Cluster 2001, as well as adding their valuable comments at various stages during the events organization. As you will agree they have managed to put together a great technical program. In addition, they have tempted the leading practitioners in the field to the conference to report on their cluster-related research and development efforts. We are sure that you will agree with us that the Cluster 2001 program is very exciting and we look forward hearing about a whole range of new and emerging developments in the field.

A conference would not be successful without an active and motivated group of local volunteers. We would like to thank Nalini Venkatasubramanian (UCI) and Susan Powell (Caltech) for all their help with the local arrangements. In addition we would like to thank Sarah Emery Bunn for her great work with the conference Web site, Kathy Little (JPL) for the advertising and technical program graphics design, and also all the student volunteers for their help ensuring that the day-to-day needs of the conference, speakers and delegates have been smoothly dealt with. Our thanks also go to the publications co-chairs, Rajkumar Buyya (Monash) and Marcin Paprzycki (Oklahoma State), our financial chair, Timothy Pinkston (USC), and our tutorials chair, Ira Pramanick (Sun). Their help has been invaluable.

Another feature of conference organization that often gets ignored is commercial sponsorship, an aspect that provides the subsidies for the likes of meals, receptions, and conference registration. We would like to thank Rick Stevens (ANL), Kathya Zamora (JPL) and Ivan Judson (ANL) for all their efforts to obtain exhibit and corporate sponsorship. Here we must also say a big thank you to our sponsors and exhibitors for their contributions to Cluster 2001. It is evident that without their help we could not have provided such a rich experience for all the attendees.

We would also like to acknowledge the IEEE Computer Society. Without their sponsorship and assistance Cluster 2001 would not have happened. In particular we would personally like to thank Mary-Kate Rada, for her patient and helpful interactions with us during the whole of organizational planning and set up of the event. Also, we extend our thanks to Doris Albritton for help, especially with all matters related to delegate registration.

Obviously there are to many other volunteers to name individually, but without their help the conference would not have happened. Finally, we would like to thank all the delegates who attend the conference, as it is you that will make Cluster 2001 a success.

By the way, our next event, Cluster 2002, is being held in Chicago in September 2002 and is being hosted jointly by Argonne National Laboratory and NCSA. We hope very much to see you all again in Chicago next year.

Enjoy your visit to the USA, Los Angeles and Newport Beach!

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

Larry Bergman
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA, USA

Message from the Program Chairs

Welcome to the IEEE 2001 International Conference on Cluster Computing.

This year marks both the dawn of the 21st Century and the emergence of Cluster Computing as the dominant model for scalable computing systems in this new epoch. Throughout the previous decade, contributors from the international community have devised a host of methodologies and developed an array of hardware and software tools that through their synergy are making possible new advances in total peak performance, price performance, and portability. Even now, a substantial number of the most powerful computers in the world are Commodity Clusters. In the next few years, the biggest systems being developed by both the US Department of Energy and National Science Foundation are commodity clusters. But even in the mainstream of computing, clusters are making dramatic inroads both for scientific/technical and commercial workloads and markets. This experience is being felt internationally with large commodity clusters serving the Asian and European communities as well. From the smallest Beowulf-class systems that permeate diverse disciplines, organizations, and markets to enormous performance behemoths, commodity clusters are finding roles in almost every arena of scalable computing. It is quite possible that commodity clusters have evolved to the stage that they now represent the long sought after "convergent parallel architecture" providing user confidence in the longevity of software investment across multiple successive generations of hardware systems. For the first time, the fortunes and whims of a single vendor cannot disrupt user productivity or make obsolete years of code development as other subsystem products can easily provide equivalent services. At the same time, vendors can be assured of growing customer acceptance, software base, and ISV targeting resulting in a strong and sustainable market. 2001 is definitely the year of the cluster.

The success and rapid growth of cluster computing is a consequence of many concurrent threads of research and development that have provided the foundations of functionality and capability essential to this new technology. Hardware advances have provided the fuel for the explosive growth in raw capability. Microprocessors and memory have improved by more than two orders of magnitude in performance and capacity, respectively, in less than a decade. Network bandwidth has been enhanced by at least two orders of magnitude as well while latency of communications has dropped by a factor of 10 to 100 in that time. Together with the expansion in scale of system size, peak performance of the largest clusters has achieved truly revolutionary growth of a factor of 100,000X in less than a decade; an unprecedented measure of advance in essentially any human endeavor. But cluster computing would be close to impossible if limited to only hardware advances. Software infrastructure – controlling the individual units and providing a framework for their cooperative operation on joint workloads and application programs – was the critical element that evolved throughout the last decade resulting in Linux and Windows as leading node operating systems and the MPI and PVM programming libraries as leading message passing program models. While these made possible basic parallel execution, far more work was required for system resource management, scheduling, and administration software. While much progress has been made in these areas in the most recent years, serious challenges must still be addressed. Of equal importance has been the creation of new algorithms that allow complex applications to be performed . These algorithms are latency tolerant and as such, they can operate effectively even on loosely coupled systems such as commodity clusters. As a result, clusters have found a much wider range of application than was once originally thought possible. Many more applications have been found, especially in the commercial sector, that are particularly suitable for commodity clusters; greatly increasing their potential market and driving industry investment in new

packaging techniques to improve foot-pad. Today more than 40 processors can be mounted in a single rack (and in certain cases, substantially more) where earlier schemes only permitted half this number. In spite of these dramatic advances, much important work remains to be done before commodity clusters fully satisfy the requirements of the broad computing community.

The IEEE Cluster 2001 Conference is the seminal community wide forum in the field of cluster computing, offering a comprehensive view of the state of the art in the enabling technologies and methodologies defining the present practices and future directions of cluster computing. The conference is organized as a series of multi-session technical tracks, each representing a critical area of advance. These include:

- *Networks* – important advances continue to drive bandwidths up and latencies down with new architectures such as VIA and IBA
- *System Management* – improving availability of means for administering, maintaining, and controlling distributed system resources to ease system usage and provide single system image.
- *Parallel I/O* – many technical and commercial applications are I/O bound as they rely on the manipulation of large data sets.
- *Scheduling* – improving the efficiencies and flexibility of scheduling parallel programs and job streams is crucial to enhances the effectiveness of clusters.
- *Middleware* – the many elements of the software infrastructure that harnesses the combined hardware resources and coordinates parallel execution.
- *Algorithms* – as new latency tolerant algorithms are developed, the range of applications to which clusters may be effectively employed becomes ever wider.
- *Node O/S* – while Windows and Linux dominate the cluster node O/S arena, other experimental and commercial operating systems are being used effectively.

To facilitate interactive dialogue and active collegial exchange of ideas, three panels will be presented on 1) Cluster Hardware, 2) Cluster Software, and 3) Applications. These panels include experts in the field that will convey distinct, often controversial, viewpoints and provide thought provoking insights in to the alternative concepts and approaches that may drive the future of cluster computing.

The success of Cluster 2001 is the product of many people's efforts including organizers, paper authors, reviewers, the vice chairs (Gordon Bell, Dave Culler, Jack Dongarra, Jim Gray, Bill Gropp, Ken Kennedy, Dan Reed, Chuck Seitz, and Burton Smith), the program committee, and the IEEE. To all of these, we wish to extend our warmest appreciation. We also wish to add our thanks to you for your participation.

Thomas Sterling
Program Chair

Daniel S. Katz
Deputy Program Chair

IEEE Cluster 2001 Conference Organization Committee

General Chairs

Mark Baker
University of Portsmouth, UK

Larry Bergman
NASA JPL, USA

Vice General Chairs

Rick Stevens
Argonne National Laboratory, USA

Nalini Venkatasubramanian
University of California at Irvine, USA

Finance Chair

Timothy Pinkston
University of Southern California, USA

Tutorial Chair

Ira Pramanick
Sun Microsystems, USA

Publications/Proceedings Co-Chairs

Marcin Paprzycki
Oklahoma State University, USA

Rajkumar Buyya
Monash University, Australia

Exhibition Chairs

Kathya Zamora
NASA JPL, USA

Ivan Judson
Argonne National Laboratory, USA

Publicity Chair

Hai Jin
Huazhong University of Science and Technology, China

Poster Chair

Phil Merkey
Michigan Technical University, USA

Steering Committee

Mark Baker
University of Portsmouth, UK

Pete Beckman
Turbolinux, Inc., USA

Bill Blake
Compaq, USA

Rajkumar Buyya
Monash University, Australia

Giovanni Chiola
DISI - Universita di Genova, Italy

Jack Dongarra
University of Tennessee and ORNL, USA

Geoffrey Fox
Indiana University, USA

Al Geist
ORNL, USA

Rusty Lusk
Argonne National Laboratory, USA

Paul Messina
Caltech, USA

Greg Pfister
IBM, Advanced Technology & Architecture, Server Design, USA

Wolfgang Rehm
Technische Universität Chemnitz, Germany

Thomas Sterling
Caltech & NASA JPL, USA

Rick Stevens
Argonne National Laboratory, USA

Thomas Stricker
ETH Zürich, Switzerland

Barry Wilkinson
UNCC, USA

Cluster 2001 Technical Program Committee

Technical Program Chair

Thomas Sterling
Caltech & NASA JPL

Deputy Program Chair

Daniel S. Katz
NASA JPL

Vice Program Chairs

Gordon Bell
Microsoft Research

Dave Culler
UC Berkeley

Jack Dongarra
University of Tennessee

Jim Gray
Microsoft

Bill Gropp
Argonne National Laboratory

Ken Kennedy
Rice University

Dan Reed
UIUC

Chuck Seitz
Myricom Inc.

Burton Smith
Cray Inc.

Program Committee

Greg Astfalk
Hewlett-Packard

Amy Apon
University of Arkansas

Mark Baker
University of Portsmouth

David Bader
University of New Mexico

David Bailey
LBL

Don Becker
Scyld Computer Company

Pete Beckman
Turbo Linux

Siegfried Benkner
University of Vienna

Kenneth P. Birman
Cornell University

Luc Bouge
ENS Lyon

Bernie Brooks
NIH

Rajkumar Buyya
Monash University

William J. Camp
Sandia National Lab

Bill Carlson
IDA-CCS

Steve Chapin
Syracuse University

Giovanni Chiola
University of Genoa

Toni Cortes
Universitat Politècnica de Catalunya.

Tom Cwik
NASA JPL

Hank Dietz
University of Kentucky

John Dorband
NASA Goddard Space Flight Center

Greg Follen
Glenn Research Center

Dennis Ganno
University of Indiana

Tarek El-Ghazawi
George Mason University

Remy Evard
Argonne National Laboratory

Al Geist
Oakridge National Lab

Salim Hariri
University of Arizona

Hermann Hellwagner
University of Klagenfurt

Atsushi Hori
Tsukuba Research Center

Yutaka Ishikawa
RWCP Japan

David Jackson
University of Utah

Ivan Judson
Argonne National Laboratory

Jeremy Kepner
Lincoln Laboratory

Vipin Kumar
University of Wisconsin

Walt Ligon
Clemson University

Miron Livny
University of Wisconsin

Josip Loncaric
ICASE

Isaac Lopez
Glenn Research Center

Glenn Luecke
Iowa State University

Rusty Lusk
Argonne National Laboratory

Luigi Mancini
University of Rome

Phil Merkey
Michigan Tech

Hans Meuer
University of Mannheim

Ron Minnich
Los Alamos National Laboratory

Abdullah Mojel
Ministry of Higher Education, Saudi Arabia

Phil Papadopoulos
UCSD

Rob Pennington
NCSA

Greg Pfister
IBM Server Technologies Division

Ira Pramanick
Sun Microsystems

Wolfgang Rehm
Technische Universitat Chemnitz

Alexander Reinefeld
ZIB Berlin

Rob Ross
Argonne National Laboratory

Isaac D. Scherson
University of California, Irvine

Tony Skjellum
MPI Software Technology Inc.

Larry Snyder
University of Washington

Thomas Stricker
ETH Zurich

Valerie Taylor
Northwestern University

Bernard Tourancheau
University C. Bernard

Putchong Uthayopas
Kasetsart University

Pearl Wang
George Mason University

Mateo Valero
Politecnica de Catalunya

Additional Paper Reviewers

Thomas Clune
Mike Heroux
Barbara Horner-Miller
Charles Norton
Hong Ong
Hernani Pedroso
Steve Plimpton
Éric Renault
Garry Smith
Paul Springer
Alan Stagg
Bronis de Supinski
Spencer Swift

CLUSTER2001

Table of Contents

Alternative Hardware

Applications I

Scheduling

Plenary Talks

Posters