Preface

The initial idea leading to cluster[1] computing was developed in the 1960s by IBM as a way of linking large mainframes to provide a cost-effective form of commercial parallelism. During those days, IBM's HASP (Houston Automatic Spooling Priority) system and its successor, JES (Job Entry System), provided a way of distributing work to a user-constructed mainframe cluster. IBM still supports clustering of mainframes through their Parallel Sysplex system, which allows the hardware, operating system, middleware, and system management software to provide dramatic performance and cost improvements while permitting large mainframe users to continue to run their existing applications.

However, cluster computing did not gain momentum until three trends converged in the 1980s: high performance microprocessors, high-speed networks, and standard tools for high performance distributed computing. A possible fourth trend is the increased need of computing power for computational science and commercial applications coupled with the high cost and low accessibility of traditional supercomputers. These building blocks are also known as killer-microprocessors, killer-networks, killer-tools, and killer-applications, respectively. The recent advances in these technologies and their availability as cheap and commodity components are making clusters or networks of computers (PCs, workstations, and SMPs) an appealing vehicle for cost-effective parallel computing. Clusters, built using commodity-off-the-shelf (COTS) hardware components as well as free, or commonly used, software, are playing a major role in redefining the concept of supercomputing.

The trend in parallel computing is to move away from specialized traditional supercomputing platforms, such as the Cray/SGI T3E, to cheaper, general purpose systems consisting of loosely coupled components built up from single or multiprocessor PCs or workstations. This approach has a number of advantages, including being able to build a platform for a given budget which is suitable for a large class of applications and workloads.

This book is motivated by the fact that parallel computing on a network of computers using commodity components has received increased attention recently, and noticeable progress towards usable systems has been made. A number of re-

[1] Cluster is a collection of interconnected computers working together as a single system.

searchers in academia and industry have been active in this field of research. Although research in this area is still in its early stage, promising results have been demonstrated by experimental systems built in academic and industrial laboratories. There is a need for better understanding of what cluster computing can offer, how cluster computers can be constructed, and what the impacts of clustering on high performance computing will be.

Though a significant number of research articles have been published in various conference proceedings and journals, the results are scattered in many places, are hard to obtain, and are difficult to understand, especially for beginners. This book, the first of its kind, gathers in one place the current and comprehensive technical coverage of the field and presents it in a tutorial form. The book's coverage reflects the state-of-the-art in high-level architecture, design, and development, and points out possible directions for further research and development.

Organization

The book is a collection of chapters written by leading scientists active in the area of parallel computing using networked computers. The primary purpose of the book is to provide an authoritative overview of this field's state-of-the-art. The emphasis is on the following aspects of cluster computing:

- Requirements, Issues, and Services
- System Area Networks, Communication Protocols, and High Performance I/O Techniques
- Resource Management, Scheduling, Load Balancing, and System Availability
- Possible Models for Cluster-based Parallel Systems
- Programming Models and Environments
- Algorithms and Applications of Clusters

The work on High Performance Cluster Computing appears in two volumes:

- Volume 1: Systems and Architectures
- Volume 2: Programming and Applications

This book, Volume 1, consists of 36 chapters, which are grouped into the following four parts:

- Part I: Requirements and General Issues
- Part II: Networking, Protocols, and I/O
- Part III: Process Scheduling, Load Sharing, and Balancing

- Part IV: Representative Cluster Systems

Part I focuses on cluster computing requirements and issues related to components, single system image, high performance, high availability, scalability, deployment, administration, and wide-area computing. Part II covers system area networks, light-weight communication protocols, and I/O. Part III discusses techniques and algorithms of process scheduling, migration, and load balancing along with representative systems. Part IV covers system architectures of some of the popular academic and commercial cluster-based systems such as Beowulf and SP/2.

Readership

The book is primarily written for graduate students and researchers interested in the area of parallel and distributed computing. However, it is also suitable for practitioners in industry and government laboratories.

The interdisciplinary nature of the book is likely to appeal to a wide audience. They will find this book a valuable source of information on recent advances and future directions of parallel computation using networked computers. This is the first book addressing various technological aspects of cluster computing in-depth, and we expect that the book will be an informative and useful reference in this new and fast growing research area.

The organization of this book makes it particularly useful for graduate courses. It can be used as a text for a research-oriented or seminar-based advanced graduate course. Graduate students will find the material covered by this book to be stimulating and inspiring. Using this book, they can identify interesting and important research topics for their Master's and Ph.D. work. It can also serve as a supplementary book for regular courses, taught in Computer Science, Computer Engineering, Electrical Engineering, and Computational Science and Informatics Departments, including:

- Advanced Computer Architecture
- Advances in Networking Technologies
- High Performance Distributed Computing
- Distributed and Concurrent Systems
- High Performance Computing
- Parallel Computing
- Networked Computing
- Trends in Distributed Operating Systems
- Cluster Computing and their Architecture

Cluster Computing Resources on the Web

The various software systems discussed in this book are freely available for download through the Internet. Please visit this book website,

- http://www.phptr.com/ptrbooks/ptr_0130137847.html

for pointers/links to further information on downloading Educational Resources, Cluster Computing Environments, and Cluster Management Systems.

Acknowledgments

First and foremost, I am grateful to all the contributing authors for their time, effort, and understanding during the preparation of the book.

I thank Albert Zomaya (University of Western Australia) for his advice and encouragement while starting this book project.

I would like to thank Kennith Birman (Cornell University), Marcin Paprzycki (University of Southern Mississippi), Hamid R. Arabnia (The University of Georgia), and Gregory F. Pfister (IBM Server Group) for their critical comments and suggestions on improving the book.

I thank Toni Cortes (Universitat Politecnica de Catalunya) for his consistent support and invaluable LaTeX expertise.

I thank Mark Baker (University of Portsmouth), Erich Schikuta (Universitaet Wien), Dror G. Feitelson (Hebrew University of Jerusalem), Daniel F. Savarese and Thomas Sterling (California Institute of Technology), Ira Pramanick (Silicon Graphics Inc), and Daniel S. Katz (Jet Propulsion Laboratory, California Institute of Technology) for writing an overview for various parts of the book.

I thank my wife, Smrithi, and my daughter, Soumya, for their love and understanding (my long absences from home) during the preparation of the book.

I acknowledge the support of the Australian Government Overseas Postgraduate Research Scholarship, the Queensland University of Technology Postgraduate Research Award (Programming Languages and Systems Research Centre Scholarship), and the Monash University Graduate Scholarship.

I thank Clemens Szyperski (Queensland University of Technology) and David Abramson (Monash University) for advising my Ph.D research program.

Finally, I would like to thank the staff at Prentice Hall, particularly Greg Doench, Mary Treacy, Joan McNamara, Barbara Cotton, and Mary Loudin. They were wonderful to work with!

Rajkumar Buyya
Monash University, Melbourne, Australia
(rajkumar@dgs.monash.edu.au / buyya@bigfoot.com)

February, 1999