# Foundations of Computer Systems Research

## 计算机系统研究基础

（英文版）

施巍松

高等教育出版社
HIGHER EDUCATION PRESS

# Foundations of Computer Systems Research

## 计算机系统研究基础(英文版)

施巍松

高等教育出版社·北京
HIGHER EDUCATION PRESS BEIJING

**图书在版编目(CIP)数据**

计算机系统研究基础 = Foundations of Computer Systems Research : 英文 / 施巍松主编. — 北京 : 高等教育出版社, 2010.10

ISBN 978-7-04-029063-9

Ⅰ. ①计… Ⅱ. ①施… Ⅲ. ①计算机系统-英文 Ⅳ. ①TP30

中国版本图书馆CIP数据核字(2010)第195231号

**策划编辑** 刘 英 **责任编辑** 刘 英 **封面设计** 张 楠 **责任印制** 陈伟光

| | | | |
|---|---|---|---|
| **出版发行** | 高等教育出版社 | **购书热线** | 010-58581118 |
| **社 址** | 北京市西城区德外大街4号 | **咨询电话** | 400-810-0598 |
| **邮政编码** | 100120 | **网 址** | http://www.hep.edu.cn |
| | | | http://www.hep.com.cn |
| **经 销** | 蓝色畅想图书发行有限公司 | **网上订购** | http://www.landraco.com |
| **印 刷** | 涿州市星河印刷有限公司 | | http://www.landraco.com.cn |
| | | **畅想教育** | http://www.widedu.com |
| **开 本** | 787×1092 1/16 | **版 次** | 2010年10月第1版 |
| **印 张** | 17.75 | **印 次** | 2010年10月第1次印刷 |
| **字 数** | 430 000 | **定 价** | 59.00元 |

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。



**物料号** 29063-00

*To my parents, my wife Wei,*
*and my daughters, Ivy and Macy*

# Preface

The field of computer systems is the core of computer science, and is concerned with how to design and implement a "good" system that is able to satisfy a variety of requirements and needs from end users and applications. The definition of "good" depends on to whom we are talking, and it could mean one or several of these basic requirements, to name a few: reliability, scalability, availability, usability, adaptability, agility, dependability, and performability. We envision that more and more new requirements will emerge as computing becomes more and more transparent and embedded in our daily life. This unstoppable trend, however, brings great challenge to future computer systems designers and practitioners. There is a gap between what the students learned in the classroom and what they are going to build in the real world. The goal of this book is to present the fundamental knowledge and techniques by laying out the foundations for the students who want to become researchers in the general area of computer systems. The book is intended for both senior undergraduate students and junior graduate students in the fields of computer science and computer engineering. Practitioners and systems designers in industry and research laboratories will find the book a very useful reference. One important principle guiding the writing of this book is that it should contain the material I would want my own students to learn before beginning their research.

The focus of systems research has changed considerably since the inception of computers. However, we have witnessed that several fundamental techniques and principles have been frequently used in systems research in the last fifty years. The goal of this book is to provide beginning computer systems researchers enough background for undertaking further systems research in their career path. The philosophy of this book is *foundation*, in other words, only the techniques that will at least survive for more than 10 years are selected.

In addition to general information for computer systems research, we have selected 15 specific topics that can be applied in the three phases of systems research, including *design*, *implementation* and *evaluation*. The 15 topics are selected based on the research experience of the authors in the past. It is by no means the complete list of techniques and probably has the bias of the authors. As time goes on, we will add more topics once we think a technique is fundamental enough to be included in the book. You are more than welcome to contribute new topics to this set by contacting the author directly by emailing at the following address `csr.foundations@gmail.com`. Supplemental materials of this book will be available at `http://www.cs.wayne.edu/~weisong`.

These 17 chapters of the book are organized into four parts, *General*, *Design*, *Implementation* and *Evaluation*, listed as follows:

1. Part I: General
   - Elements
   - Rules of Thumb

2. Part II: Design
   - Bloom Filters
   - Distributed Hash Tables
   - Locality Sensitive Hashing
   - XOR Operations
   - Adaptation
   - Optimistic Replication
   - Reputation and Trust
   - Moving Average
   - Machine Learning
3. Part III: Implementation
   - Asynchronous I/O
   - Multithreading
   - Virtualization
4. Part IV: Evaluation
   - Queueing Theory
   - Black Box Testing
   - Goodness-of-Fit

Following the first two chapters about the elements and rules of thumb of systems research, each topic will be covered by one chapter. To help readers understand these topics easily, we also provide two case studies of each topic at the end of each chapter (if possible), in addition to the basic description of the motivation and idea. These case studies are selected from different aspects of computer systems design and implementation.

## Acknowledgement

This book would not be possible without the interactions with my past and current students at the Mobile and Internet Systems Laboratory at Wayne State University and several collaborators in China. I am greatly indebted to them: Hanping Lufei, Zhengqiang Liang, Kewei Sha, Safwan Al-Omari, Zhifeng Yu, Jayashree Ravi, Sharun Santhosh, Yonggeng Mao, Brandon Szeliga, Yong Xi, Tung Nguyen, Guoxing Zhan, Suhib Rawshdeh, Shinan Wang, Thanh Do Duc, Chenjia Wang, Kevin Monaghan, John Cavicchio, Hui Chen, Lingjun Fan, Huajia Mao, Ying Song, Yaqiong Li, Jianfeng Zhan, and Yongqiang He. The interactions with them directly motivated me to write this book. I also would like to thank the students who attended the distributed systems course I taught at the Institute of Computing Technology (ICT), Chinese Academy of Sciences, and the 2009 Dragon Star course entitled "Principles of Computer Systems Design" I taught at Tsinghua University in May 2009. The Dragon Star program was partially sponsored by the National Natural Science Foundation of China and ICT. Early feedbacks from them greatly helped reshape the content of this book. As a matter of fact, the majority of the book's contents were written by the volunteer students from these two courses, under my guidance. Without their hard work, we would have not seen this print for at least another couple of years.

I am extremely fortunate to have the opportunity to work with the authors of Chapter 3 to Chapter 17 of the book, including Dajun Lu, Yanbin Liu, Wei Jiang, Dawen Xu, Mingdong Tang, Wei Lee, Dandan Tu, Mingwen Chen, Shinan Wang, Nan Yuan, Yongbin Zhou, Changlin Wan, Tung Nguyen, Yanan Li, Zhuhua Liao, Zhengqiang Liang, and Yong Zhao. All of them are dedicated and enthusiastic on their corresponding chapters. Without their tremendous effort, you would have not been able to see the book presented in such a coherent way.

I want to thank Professor Zhiwei Xu, the Chief Technology Officer of Institute of Computing

Technology, Chinese Academy of Sciences, who invited me to spend six months at ICT during my sabbatical leave. My appreciation also goes to ICT, including the Graduate Education Office and their staff, Key Laboratory of Network Science and Technology, Key Laboratory of Systems and Architecture, and so on. Without their support, I would not have had a chance to meet and interact with fellow students at ICT. It is these interactions that inspired me to complete the text as soon as possible. I would like to thank Professor Wenguang Chen from Tsinghua University, Professor Nong Xiao from National University of Defense Technology and Professor Limin Sun from Institute of Software, Chinese Academy of Sciences, who gave me a great opportunity to teach an early version of this book at their respective institutions. I also would like to thank many friends, especially Professor Yafei Dai from Peking University, Professor Yunquan Zhang from Institute of Software, Chinese Academy of Sciences, for their early interests to the book convinced me that writing this book is the right direction.

My deepest appreciation also goes to Ms. Ying Liu from Higher Education Press. It is her great enthusiasm and support that made this book be published so smoothly. I want to thank Scott D. Vance, Jacqueline D. Brown, who gave great comments on improving the presentation of this book.

I also would like to thank the support from Wayne State University and the colleagues at the Department of Computer Science. The generous support for my sabbatical leave (Winter 2009) and one course release as part of the 2009—2010 Wayne State University Career Development Chair Award (2009—2010) which afforded me enough time to focus on writing this book.

Last but not least, I am greatly indebted to my family: my parents, my wife, and my daughters. Without my parents' help with baby caring, I would have not been able to dedicate too much time to the book this year. My wife, as always, supported me in writing this book and pushed me to complete it as early as possible. Without her great support and understanding, I don't think this publication would have been possible. I also want to thank my two daughters, Ivy and Macy, who showed great understanding and lent cooperation. They are responsible for this book in more ways than even they realize. This book is dedicated to them.

Enjoy the book!

*Weisong Shi*
Detroit
August 2010

# Contents

## Part I General

## Part II Design

# Part I
# General

Part I consists of two chapters. Chapter 1 describes the basic elements of computer systems research, including most common mistakes in English writing, and Chapter 2 lists 12 rules of thumb that are widely used in systems design.

# Chapter 1
# Elements

Weisong Shi

**Abstract** The content of this chapter lays the foundations for computer systems researchers. If you are an experienced researcher in this area, you can skip the chapter and move directly to the next chapter. The first thing for a beginner is to identify the top conferences/journals in his or her specific research area. Hence, the 20 top conferences and the 10 top journals in the computer systems area are listed in the first part of the chapter. You might want to customize them a little bit to fit your own interests. Next, three "how tos" are described respectively: *how to read a research paper*, *how to write a research paper*, and *how to present a research paper*. Finally, the chapter is concluded by pointing interested readers to an important book about the ethical foundations of scientific practices.

## 1.1 Top Systems Conferences/Journals

In this section, the 20 top systems or systems-related conferences and the 10 top journals in computer systems are listed based on the author's personal research experience; and it is by no means a complete list. Note that among these 20 conferences, the first half is more systems-oriented, while the second half has its own specific interests, although they do have a couple of sections focusing on systems-related topics. Note that the list intends to give beginners a point to start, rather than an authoritative list. As a matter of fact, you are encouraged to work out a customized list based on your own research interests.

It is worth noting that at the time this book was written (2009), top conferences were more prestigious than journals in the computer systems community. Therefore, the 20 top systems conferences are listed first. In the long run, however, the systems community is called on to pay more attention to journals than conferences in the future because of the following two reasons. One is that the authors of a conference paper have little time (usually less than a month between the notification date and the camera ready date) to revise the paper according to review comments. Therefore, most of results published in conference proceedings are premature work. On the contrary, journal papers are archived documents, and authors have enough time to revise the paper and address the comments. The second reason is related to the development of computer science as a discipline. To the best of my knowledge, many other disciplines pay more attention to journals other than conferences. The differences in evaluation (i.e., conferences vs. journals) that exist between computer

Weisong Shi
Wayne State University, Detroit, MI 48202, USA. e-mail: weisong@wayne.edu