

Lecture Notes in Computer Science

Edited by G. Goos and J. Hartmanis

448

B. Simons A. Spector (Eds.)

Fault-Tolerant
Distributed Computing



Springer-Verlag

New York Berlin Heidelberg London

Paris Tokyo Hong Kong Barcelona

Editorial Board

D. Barstow W. Brauer P. Brinch Hansen D. Gries D. Luckham
C. Moler A. Pnueli G. Seegmüller J. Stoer N. Wirth

Editors

Barbara Simons

IBM Almaden Research Center, Dept. K53/802
650 Harry Road, San Jose, CA 95120-6099, USA

Alfred Spector

Transarc Corporation, The Gulf Tower
707 Grant Street, Pittsburgh, PA 15219, USA

CR Subject Classification (1987): D.4, C.2.4, F.1.1, F.2.0, C.3–4

ISBN 3-540-97385-0 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-97385-0 Springer-Verlag New York Berlin Heidelberg

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 other ways, and storage in data banks. Duplication of this publication or parts thereof is only permitted under the provisions of the German Copyright Law of September 9, 1965, in its current version, and a copyright fee must always be paid. Violations fall under the prosecution act of the German Copyright Law.

© Springer-Verlag Berlin Heidelberg 1990
Printed in Germany

Printing and binding: Druckhaus Beltz, Hemsbach/Bergstr.
2145/3140-543210 – Printed on acid-free paper

Preface

The goal of the Asilomar Workshop on Fault-Tolerant Distributed Computing was to facilitate interaction between theoreticians and practitioners. To achieve this goal, speakers were invited and topics chosen so that a broad overview of the field would be presented. Because the attendance at the Workshop was diverse, the presentations/papers were in many instances designed to appeal to a general audience. The presentations were also designed to span a body of research from the theoretical to the pragmatic. Since the material seemed ideal for a book, this book was planned together with the planning of the workshop.

Held in the spring of 1986, the workshop brought together approximately 70 active researchers from academia and industry. Most of the chapters were written following the workshop and subsequently revised. Consequently, some of the results that are described were obtained after the workshop. However, six of the chapters (those by Bernstein, Cohn, Finkelstein, Liskov, Spector, and Wensley) were recorded, transcribed, and edited for inclusion. This transcription format makes for easy-to-read, though somewhat chatty, articles. The chapters are arranged in the order of presentation at the workshop.

Given the scope of this book, we feel that it should be of use to students, researchers, and developers. We hope that it will promote greater understanding within the world of fault-tolerant research and development.

Barbara Simons
Alfred Spector

Workshop Program Committee: Chair: Barbara Simons, Flaviu Cristian, Danny Dolev, Michael Fischer, Jim Gray, Leslie Lamport, Nancy Lynch, Marshall Pease, Fred Schneider, Ray Strong.

We would like to thank IBM Almaden Research Center and the Office of Naval Research for their support. We also appreciate the assistance of several volunteers who assisted at the workshop.

Contents

A Theoretician's View of Fault Tolerant Distributed Computing	1
<i>M. J. Fischer</i>	
A Comparison of the Byzantine Agreement Problem and the Transaction Commit Problem	10
<i>J. Gray</i>	
The State Machine Approach: A Tutorial	18
<i>F. B. Schneider</i>	
A Simple Model for Agreement in Distributed Systems	42
<i>D. Dolev, R. Strong</i>	
Atomic Broadcast in a Real-Time Environment	51
<i>F. Cristian, D. Dolev, R. Strong, H. Aghili</i>	
Randomized Agreement Protocols	72
<i>M. Ben-Or</i>	
An Overview of Clock Synchronization	84
<i>B. Simons, J. L. Welch, N. Lynch</i>	
Implementation Issues in Clock Synchronization	97
<i>M. Beck, T. K. Srikanth, S. Toueg</i>	
 Systems Session I	
Argus	108
<i>B. Liskov</i>	
TABS	115
<i>A. Z. Spector</i>	
Communication Support for Reliable Distributed Computing	124
<i>K. P. Birman, T. A. Joseph</i>	
Algorithms and System Design in the Highly Available Systems Project	138
<i>S. J. Finkelstein</i>	

Easy Impossibility Proofs for Distributed Consensus Problems	147
<i>M. J. Fischer, N. A. Lynch, M. Merritt</i>	
An Efficient, Fault-Tolerant Protocol for Replicated Data Management	171
<i>D. Skeen, A. El Abbadi, F. Cristian</i>	

Systems Session II

Arpanet Routing	192
<i>S. Cohn</i>	
On the Relationship Between the Atomic Commitment and Consensus Problems	201
<i>V. Hadzilacos</i>	
The August System	209
<i>J. Wensley</i>	
The Sequoia System	217
<i>P. Bernstein</i>	
Fault Tolerance in Distributed UNIX	224
<i>A. Borg, W. Blau, W. Oberle, W. Graetsch</i>	
Faults and Their Manifestation	244
<i>D. P. Siewiorek</i>	
The "Engineering" of Fault-Tolerant Distributed Computing Systems	262
<i>Ö. Babaoğlu</i>	
Bibliography for Fault-Tolerant Distributed Computing	274
<i>B. A. Coan</i>	