

# Lecture Notes in Computer Science

938

Edited by G. Goos, J. Hartmanis and J. van Leeuwen

Advisory Board: W. Brauer D. Gries J. Stoer

K.P. Birman F. Mattern A. Schiper (Eds.)

# Theory and Practice in Distributed Systems

International Workshop  
Dagstuhl Castle, Germany, September 5-9, 1994  
Selected Papers



Springer

## Series Editors

Gerhard Goos

Universität Karlsruhe

Vincenz-Priessnitz-Straße 3, D-76128 Karlsruhe, Germany

Juris Hartmanis

Department of Computer Science, Cornell University

4130 Upson Hall, Ithaca, NY 14853, USA

Jan van Leeuwen

Department of Computer Science, Utrecht University

Padualaan 14, 3584 CH Utrecht, The Netherlands

## Volume Editors

Kenneth P. Birman

Department of Computer Science, Cornell University

Ithaca, NY 14853, USA

Friedemann Mattern

FB 20 Informatik, Technische Hochschule Darmstadt

Alexanderstr. 10, D-64283 Darmstadt, Germany

André Schiper

Department of Computer Science

Swiss Federal Institute of Technology Lausanne

CH-1015 Lausanne, Switzerland

Library of Congress Catalog-in-Publication Data. Theory and practice in distributed systems: international workshop. Dagstuhl Castle, Germany, September 5-9, 1994: selected papers/K. P. Birman, F. Mattern, A. Schiper, eds. p.cm. - (Lecture notes in computer science; 938). Includes bibliographical references and index. ISBN 3-540-60042-6 (Berlin: acid-free paper). - ISBN 0-387-60042-6 (New York: acid-free paper)  
1. Electronic data processing--Distributed processing--Congresses. I Birman, Kenneth P. II. Mattern, Friedemann. III. Schiper, André, 1949-. IV. series, QA76.9.D5T48 1995. 004'.36--dc20 95-21873 CIP

CR Subject Classification (1991): C.2.4

ISBN 3-540-60042-6 Springer-Verlag Berlin Heidelberg New York

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 any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable for prosecution under the German Copyright Law.

© Springer-Verlag Berlin Heidelberg 1995  
Printed in Germany

Typesetting: Camera-ready by author  
SPIN: 10486321 06/3142 - 543210 - Printed on acid-free paper

This volume is dedicated to Flaviu Cristian, who was unable to attend the workshop due to health problems.

His assistance in developing the concept of the workshop, organizing it, and selecting the invited participants is gratefully acknowledged.

## Preface

During the past 20 years, a substantial theoretical and practical base has evolved in the area of distributed computing. However, this work has been done by largely disjoint sets of researchers, with the result that much theory is inapplicable to real-world systems, and many of the real-world systems that have been built suffer from weaknesses that could be overcome using the existing theoretical methodology. It was therefore the intention of the workshop "Unifying Theory and Practice in Distributed Systems", held from September 5th to September 9th, 1994 at Schloß Dagstuhl, to bring together a diverse group of pragmatically inclined theoreticians and theoretically inclined practitioners. The goal of the workshop was to share insight, educate one another, and lay the groundwork for the next generation of distributed systems research and development.

The workshop was organized by Kenneth Birman (Cornell University), Flaviu Cristian (University of California at San Diego), Friedemann Mattern (University of Saarland at Saarbrücken), and André Schiper (Ecole Polytechnique de Lausanne). It brought together established experts from academia and industry as well as young scientists from nine countries. In total, 30 talks (some with on-line and video demonstrations) were given during the week, often followed by lively and sometimes controversial discussions. In two extra evening discussion sessions the implications of the Fischer-Lynch-Paterson Theorem (impossibility of distributed consensus with one faulty processor) for practical system design and the various notions of "real time" were explored.

The papers in this volume were revised from position papers presented by the lecturers, and have been reviewed by two referees. Final preparation and revision occurred after the workshop, so that the papers might reflect insights gained through discussions at the workshop. The papers place emphasis on the following general themes:

- important paradigms and influential concepts;
- fundamental algorithms and principles;
- fault-tolerance;
- real-time;
- system structures, basic services, toolkits;
- large scale issues, application domains, case studies.

In two extra evening sessions during the workshop, the implications of the Fischer-Lynch-Paterson impossibility result for practical system design, and the various notions of "real-time" were explored. The discussions are summarized at the end of the proceedings.

In total, the seminar was considered to be successful and very interesting. The special nature and atmosphere of Schloß Dagstuhl offered ample opportunities for personal discussions, the computer science library was also used extensively by some participants. The fine food, the good wine, and the perfect organization by the office and the local staff of Schloß Dagstuhl were much appreciated by all participants. The organizers also gratefully acknowledge the support of the

Dagstuhl International Conference and Research Center for Computer Science and thank Springer-Verlag for publishing the proceedings in the LNCS series.

It is hoped that through these proceedings, interested individuals who were not able to attend can share with us the debates that characterized the workshop, and join with us in reviewing important advances in understanding the role of theory in distributed systems research!

April 1995

Kenneth Birman  
Friedemann Mattern  
André Schiper

# Table of Contents

## Experimental systems

Architectural issues in the StormCast system <i>D. Johansen, G. Hartvigsen</i> .....	1
---	---

Lessons learned from building and using the <i>Arjuna</i> distributed programming system <i>S. K. Shrivastava</i> .....	17
--	----

## Group communication

A high performance totally ordered multicast protocol <i>B. Whetten, T. Montgomery, S. Kaplan</i> .....	33
--	----

New applications for group computing <i>R. van Renesse, K. P. Birman, T. von Eicken, K. Marzullo</i> .....	58
---	----

Support for information sharing in CSCW based on causally and totally ordered group communication <i>M. Kolland</i> .....	64
--	----

The design of the Transis system <i>D. Dolev, D. Malki</i> .....	83
---	----

The Rampart toolkit for building high-integrity services <i>M. K. Reiter</i> .....	99
---	----

## Fault-tolerance issues in distributed systems

Deriving optimal checkpoint protocols for distributed shared memory architectures <i>L. Alvisi, K. Marzullo</i> .....	111
--	-----

Transaction model vs virtual synchrony model: bridging the gap <i>R. Guerraoui, A. Schiper</i> .....	121
---	-----

## System structuring and structuring tools

Modelling Darwin in the $\pi$ -calculus <i>J. Magee, S. Eisenbach, J. Kramer</i> .....	133
Towards open service environments <i>K. Geihs, B. Bär, A. Puder</i> .....	153

## Correctness issues

Correctness proofs of distributed algorithms <i>W. Reisig</i> .....	164
Deterministic fault injection of distributed systems <i>S. Dawson, F. Jahanian</i> .....	178

## Consistency issues and ordering algorithms

A non-blocking lightweight implementation of causal order message delivery <i>F. Mattern, S. Fünfrocken</i> .....	197
Merits of a probabilistic approach to properties in process group systems <i>K. P. Birman</i> .....	214
Sequential consistency in distributed systems <i>M. Mizuno, M. Raynal, J. Z. Zhou</i> .....	224

## Parallelism

Speedup limits for tightly-coupled parallel computations <i>R. Schwarz</i> .....	242
Summary of the discussion sessions: FLP and real-time .....	260

List of Authors .....	263
-----------------------	-----