

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison

Lancaster University, UK

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Alfred Kobsa

University of California, Irvine, CA, USA

Friedemann Mattern

ETH Zurich, Switzerland

John C. Mitchell

Stanford University, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz

University of Bern, Switzerland

C. Pandu Rangan

Indian Institute of Technology, Madras, India

Bernhard Steffen

TU Dortmund University, Germany

Madhu Sudan

Microsoft Research, Cambridge, MA, USA

Demetri Terzopoulos

University of California, Los Angeles, CA, USA

Doug Tygar

University of California, Berkeley, CA, USA

Gerhard Weikum

Max Planck Institute for Informatics, Saarbruecken, Germany

Roberto Baldoni Nicolas Nisse
Maarten van Steen (Eds.)

Principles of Distributed Systems

17th International Conference, OPODIS 2013
Nice, France, December 16-18, 2013
Proceedings



Springer

Volume Editors

Roberto Baldoni

Sapienza Research Center of Cyber Intelligence and Information Security
and

Università degli Studi di Roma "La Sapienza"

Dipartimento di Ingegneria Informatica, Automatica e Gestionale "Antonio Ruberti"

Via Ariosto 25, 00185 Rome, Italy

E-mail: baldoni@dis.uniroma1.it

Nicolas Nisse

Inria, France

and

Université Nice Sophia Antipolis CNRS, 13S, UMR 7271

06900 Sophia Antipolis, France

E-mail: nicolas.nisse@inria.fr

Maarten van Steen

Vrije Universiteit Amsterdam

Department of Computer Science

De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands

E-mail: steen@cs.vu.nl

ISSN 0302-9743

e-ISSN 1611-3349

ISBN 978-3-319-03849-0

e-ISBN 978-3-319-03850-6

DOI 10.1007/978-3-319-03850-6

Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013954562

CR Subject Classification (1998): C.2.4, C.2, F.2, D.2, I.2.11, G.2.2

LNCS Sublibrary: SL 1 – Theoretical Computer Science and General Issues

© Springer International Publishing Switzerland 2013

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

This volume contains the papers presented at OPODIS 2013. OPODIS, the International Conference on Principles of Distributed Systems, is an international forum for the exchange of state-of-the-art knowledge on distributed computing and systems among researchers from around the world. The 17th edition of OPODIS was held during December 16–18, 2013 in Nice, France.

Papers were sought soliciting original research contributions to the theory, specification, design and implementation of distributed systems. In response to the call for papers, 41 submissions were received, out of which 18 papers were accepted, after a rigorous reviewing process that involved 33 Program Committee members and at least three reviews per paper.

We would like to thank the Program Committee members, as well as the external reviewers, for their fundamental contribution in selecting the best papers.

In addition to the technical papers, the program included five invited presentations by: Marcos k. Aguilera (Microsoft Research, USA), Eitan Altman (Inria, France), Hein Meling (University of Stravanger, Norway), Nuno Preguica (Universidade Nova de Lisboa, Portugal) and Marc Shapiro (Inria, France).

September 2013

Roberto Baldoni
Nicolas Nisse
Maarten van Steen

Organization

Program Committee

| | |
|------------------------|---|
| Marco Aiello | University of Groningen, The Netherlands |
| Roberto Baldoni | Università di Roma “La Sapienza”, Italy |
| Christian Cachin | IBM Research, Zurich, Switzerland |
| Antonio Carzaniga | University of Lugano, Switzerland |
| Gregory Chockler | IBM Haifa Research Laboratory, Israel |
| Allen Clement | Max Planck Institute for Software Systems, Germany |
| Paolo Costa | Microsoft Research Cambridge, UK |
| Dick Epema | Delft University of Technology, The Netherlands |
| Patrick Eugster | Purdue University, USA |
| Pascal Felber | Université de Neuchâtel, Switzerland |
| Antonio Fernandez Anta | Institute IMDEA Networks, Spain |
| Paola Flocchini | University of Ottawa, Canada |
| Ali Ghodsi | University of California at Berkeley, USA |
| Rachid Guerraoui | EPFL, Switzerland |
| Aaron Harwood | University of Melbourne, Australia |
| Konrad Iwanicki | University of Warsaw, Poland |
| Mark Jelasity | University of Szeged, Hungary |
| Ricardo Jimenez Peris | Universidad Politécnica de Madrid, Spain |
| Anne-Marie Kermarrec | Inria, France |
| Hein Meling | University of Stavanger, Norway |
| Alessia Milani | Bordeaux Institute of Technology, France |
| Alberto Montresor | University of Trento, Italy |
| Nicolas Nisse | Inria, France |
| Peter Pietzuch | Imperial College London, UK |
| Maria Potop-Butucaru | Université Pierre et Marie Curie-LIP6, France |
| Luis Rodrigues | Universidade de Lisboa, Portugal |
| Cristina Seceleanu | Malardalen University, Sweden |
| Marc Shapiro | Inria, Univ. Pierre et Marie Curie-LIP6, France |
| Alex Shraer | Google, USA |
| Peter Triantafillou | University of Glasgow, Scotland |
| Frits Vaandrager | Radboud University Nijmegen, The Netherlands |
| Maarten van Steen | VU University Amsterdam, The Netherlands |
| Paulo Verissimo | Universidade de Lisboa, Portugal |

Roman Vitenberg
Spyros Voulgaris
Masafumi Yamashita

University of Oslo, Norway
Vrije Universiteit Amsterdam, The Netherlands
Kyushu University, Fukuoka, Japan

Additional Reviewers

Anagnostopoulos, Christos
Ananthanarayanan, Ganesh
Arad, Cosmin
Culhane, William
Dobre, Dan
Georgievski, Ilche
Godard, Emmanuel
Jehl, Leander
Kogan, Kirill
Lamani, Anissa
Lehmann, Anja
Marinescu, Raluca
Mostefaoui, Achour
Pagani, Giuliano Andrea

Perelman, Dmitri
Schmid, Ulrich
Sens, Pierre
Shafaat, Tallat M.
Shavit, Nir
Sutra, Pierre
Taherkordi, Amir
Travers, Corentin
Trehan, Amitabh
Tretmans, Jan
Tso, Posco
Urdaneta, Guido
Verbeek, Freek
Vilaca, Xavier

Invited Talks

Geo-Distributed Storage in Data Centers

Marcos K. Aguilera

Microsoft Research
Mountain View, CA, USA

Abstract. Data centers increasingly have a storage system that is *geo-distributed*, that is, distributed across several geographic locations. We explain the general characteristics of this setting and the challenges that it brings, chief among them the need to operate with low latency despite significant network delays. These challenges lead to many interesting problems: migrating data online, dealing with congestion, providing efficient transactions, and more. We discuss these problems and some recent solutions, which bring together techniques from distributed computing, distributed systems, and database systems. Despite much progress, however, several algorithmic and fundamental questions remain open and serve as inspiration for further investigation.

Dynamic Game Models in Complex Systems

Eitan Altman *

Inria, France

`eitan.altman@inria.fr`

Abstract. We begin the tutorial with a theoretic part that covers two areas: non-cooperative game theory, and population propagation models. In the game theory part, a particular attention will be given to potential games. We shall focus in particular on congestion games and on the game version of the generalized Kelly mechanism problem, both of which are known to be potential games. In our presentation of models for population propagation models, we shall present several models which we shall classify according to the size of population of potential interested destination nodes (which can be finite and constant, finite but non-constant or infinite), and the virality of the content. This will include branching and epidemic models. We shall then use these tools to study various applications to large networks. This will include (1) security issues related to e-virus attacks, (2) the question of what type of content should service providers specialize in, which will be solved by transforming it into an equivalent congestion game, (3) issues related to viral marketing and competition issues in social networks. In these problems the generalized Kelly mechanism will be frequently used. The game theoretic analysis will allow us to get insight on how much to spend on advertising products and on what product should we advertise. Both journal and conference papers as well as video presentations covering this tutorial are available at <http://www-sop.inria.fr/members/Eitan.Altman/dodescaden.html>

* This work was supported by CONGAS Project (FP7- ICT-2011-8-317672), see www.congas-project.eu

Table of Contents

| | |
|--|-----|
| Tutorial Summary: Paxos Explained from Scratch | 1 |
| <i>Hein Meling and Leander Jehl</i> | |
| On Two-Party Communication through Dynamic Networks | 11 |
| <i>Sebastian Abshoff, Markus Benter, Manuel Malatyali, and Friedhelm Meyer auf der Heide</i> | |
| Eventual Leader Election in Evolving Mobile Networks | 23 |
| <i>Luciana Arantes, Fabíola Greve, Pierre Sens, and Véronique Simon</i> | |
| Self-stabilizing Leader Election in Population Protocols over Arbitrary Communication Graphs | 38 |
| <i>Joffroy Beauquier, Peva Blanchard, and Janna Burman</i> | |
| α -Register | 53 |
| <i>David Bonnin and Corentin Travers</i> | |
| How (Not) to Shoot in Your Foot with SDN Local Fast Failover: A Load-Connectivity Tradeoff | 68 |
| <i>Michael Borokhovich and Stefan Schmid</i> | |
| Message Passing or Shared Memory: Evaluating the Delegation Abstraction for Multicores | 83 |
| <i>Irina Calciu, Dave Dice, Tim Harris, Maurice Herlihy, Alex Kogan, Virendra Marathe, and Mark Moir</i> | |
| Reputation-Based Mechanisms for Evolutionary Master-Worker Computing | 98 |
| <i>Evgenia Christoforou, Antonio Fernández Anta, Chryssis Georgiou, Miguel A. Mosteiro, and Angel (Anxo) Sánchez</i> | |
| State-Driven Testing of Distributed Systems | 114 |
| <i>Domenico Cotroneo, Roberto Natella, Stefano Russo, and Fabio Scippacercola</i> | |
| Self-stabilizing Resource Discovery Algorithm | 129 |
| <i>Seda Davtyan, Kishori M. Konwar, and Alexander A. Shvartsman</i> | |
| Hybrid Distributed Consensus | 145 |
| <i>Roy Friedman, Gabriel Kliot, and Alex Kogan</i> | |

| | |
|--|------------|
| Speculative Concurrent Processing with Transactional Memory in the Actor Model | 160 |
| <i>Yaroslav Hayduk, Anita Sobe, Derin Harmanci, Patrick Marlier, and Pascal Felber</i> | |
| An Optimal Broadcast Algorithm for Content-Addressable Networks . . . | 176 |
| <i>Ludovic Henrio, Fabrice Huet, and Justine Rochas</i> | |
| On Local Fixing | 191 |
| <i>Michael König and Roger Wattenhofer</i> | |
| A Skiplist-Based Concurrent Priority Queue with Minimal Memory Contention | 206 |
| <i>Jonatan Lindén and Bengt Jonsson</i> | |
| VirtuCast: Multicast and Aggregation with In-Network Processing : An Exact Single-Commodity Algorithm | 221 |
| <i>Matthias Rost and Stefan Schmid</i> | |
| Mobile Byzantine Agreement on Arbitrary Network | 236 |
| <i>Toru Sasaki, Yukiko Yamauchi, Shuji Kijima, and Masafumi Yamashita</i> | |
| On Scheduling Algorithms for MapReduce Jobs in Heterogeneous Clouds with Budget Constraints | 251 |
| <i>Yang Wang and Wei Shi</i> | |
| Fast and Scalable Queue-Based Resource Allocation Lock on Shared-Memory Multiprocessors | 266 |
| <i>Deli Zhang, Brendan Lynch, and Damian Dechev</i> | |
| Author Index | 281 |