Festschrift

LNCS 4700

Cliff B. Jones
Zhiming Liu
Jim Woodcock (Eds.)

Formal Methods and Hybrid Real-Time Systems

Essays in Honour of Dines Bjørner and Zhou Chaochen on the Occasion of Their 70th Birthdays

Springer

Cliff B. Jones Zhiming Liu
Jim Woodcock (Eds.)

Formal Methods and Hybrid Real-Time Systems

Essays in Honour of Dines Bjørner and Zhou Chaochen on the Occasion of Their 70th Birthdays

Springer

Volume Editors

Cliff B. Jones
Newcastle University, School of Computing Science
Newcastle upon Tyne, NE1 7RU, UK
E-mail: cliff.jones@ncl.ac.uk

Zhiming Liu
United Nations University, International Institute for Software Technology
Macao, China
E-mail: z.liu@iist.unu.edu

Jim Woodcock
University of York, Department of Computer Science
Heslington, York YO10 5DD, UK
E-mail: jim@cs.york.ac.uk

The illustration appearing on the cover of this book is the work of Daniel Rozenberg (DADARA).

Library of Congress Control Number: 2007935177

CR Subject Classification (1998): D.2, D.3, C.3, F.3-4, C.2, H.4

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

ISSN 0302-9743
ISBN-10 3-540-75220-X Springer Berlin Heidelberg New York
ISBN-13 978-3-540-75220-2 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springer.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 12164615 06/3180 5 4 3 2 1 0

Lecture Notes in Computer Science 4700

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Lecture Notes in Computer Science

Sublibrary 1: Theoretical Computer Science and General Issues

For information about Vols. 1–4445
please contact your bookseller or Springer

Vol. 4770: V.G. Ganzha, E.W. Mayr, E.V. Vorozhtsov (Eds.), Computer Algebra in Scientific Computing. XIII, 460 pages. 2007.

Vol. 4743: P. Thulasiraman, X. He, T.L. Xu, M.K. Denko, R.K. Thulasiram, L.T. Yang (Eds.), Frontiers of High Performance Computing and Networking ISPA 2007 Workshops. XXIX, 536 pages. 2007.

Vol. 4742: I. Stojmenovic, R.K. Thulasiram, L.T. Yang, W. Jia, M. Guo, R.F. de Mello (Eds.), Parallel and Distributed Processing and Applications. XX, 995 pages. 2007.

Vol. 4736: S. Winter, M. Duckham, L. Kulik, B. Kuipers (Eds.), Spatial Information Theory. XV, 455 pages. 2007.

Vol. 4732: K. Schneider, J. Brandt (Eds.), Theorem Proving in Higher Order Logics. IX, 401 pages. 2007.

Vol. 4731: A. Pelc (Ed.), Distributed Computing. XVI, 510 pages. 2007.

Vol. 4710: C.W. George, Z. Liu, J. Woodcock (Eds.), Domain Modeling and the Duration Calculus. XI, 237 pages. 2007.

Vol. 4708: L. Kučera, A. Kučera (Eds.), Mathematical Foundations of Computer Science 2007. XVIII, 764 pages. 2007.

Vol. 4707: O. Gervasi, M.L. Gavrilova (Eds.), Computational Science and Its Applications – ICCSA 2007, Part III. XXIV, 1205 pages. 2007.

Vol. 4706: O. Gervasi, M.L. Gavrilova (Eds.), Computational Science and Its Applications – ICCSA 2007, Part II. XXIII, 1129 pages. 2007.

Vol. 4705: O. Gervasi, M.L. Gavrilova (Eds.), Computational Science and Its Applications – ICCSA 2007, Part I. XLIV, 1169 pages. 2007.

Vol. 4703: L. Caires, V.T. Vasconcelos (Eds.), CONCUR 2007 – Concurrency Theory. XIII, 507 pages. 2007.

Vol. 4700: C.B. Jones, Z. Liu, J. Woodcock (Eds.), Formal Methods and Hybrid Real-Time Systems. XVI, 539 pages. 2007.

Vol. 4697: L. Choi, Y. Paek, S. Cho (Eds.), Advances in Computer Systems Architecture. XIII, 400 pages. 2007.

Vol. 4688: K. Li, M. Fei, G.W. Irwin, S. Ma (Eds.), Bio-Inspired Computational Intelligence and Applications. XIX, 805 pages. 2007.

Vol. 4684: L. Kang, Y. Liu, S. Zeng (Eds.), Evolvable Systems: From Biology to Hardware. XIV, 446 pages. 2007.

Vol. 4683: L. Kang, Y. Liu, S. Zeng (Eds.), Intelligence Computation and Applications. XVII, 663 pages. 2007.

Vol. 4681: D.-S. Huang, L. Heutte, M. Loog (Eds.), Advanced Intelligent Computing Theories and Applications. XXVI, 1379 pages. 2007.

Vol. 4672: K. Li, C. Jesshope, H. Jin, J.-L. Gaudiot (Eds.), Network and Parallel Computing. XVIII, 558 pages. 2007.

Vol. 4671: V. Malyshkin (Ed.), Parallel Computing Technologies. XIV, 635 pages. 2007.

Vol. 4669: J.M. de Sá, L.A. Alexandre, W. Duch, D. Mandic (Eds.), Artificial Neural Networks – ICANN 2007, Part II. XXXI, 990 pages. 2007.

Vol. 4668: J.M. de Sá, L.A. Alexandre, W. Duch, D. Mandic (Eds.), Artificial Neural Networks – ICANN 2007, Part I. XXXI, 978 pages. 2007.

Vol. 4666: M.E. Davies, C.J. James, S.A. Abdallah, M.D Plumbley (Eds.), Independent Component Analysis and Blind Signal Separation. XIX, 847 pages. 2007.

Vol. 4665: J. Hromkovič, R. Královič, M. Nunkesser, P. Widmayer (Eds.), Stochastic Algorithms: Foundations and Applications. X, 167 pages. 2007.

Vol. 4664: J. Durand-Lose, M. Margenstern (Eds.), Machines, Computations, and Universality. X, 325 pages. 2007.

Vol. 4649: V. Diekert, M.V. Volkov, A. Voronkov (Eds.), Computer Science – Theory and Applications. XIII, 420 pages. 2007.

Vol. 4647: R. Martin, M. Sabin, J. Winkler (Eds.), Mathematics of Surfaces XII. IX, 509 pages. 2007.

Vol. 4646: J. Duparc, T.A. Henzinger (Eds.), Computer Science Logic. XIV, 600 pages. 2007.

Vol. 4644: N. Azémard, L. Svensson (Eds.), Integrated Circuit and System Design. XIV, 583 pages. 2007.

Vol. 4641: A.-M. Kermarrec, L. Bougé, T. Priol (Eds.), Euro-Par 2007 Parallel Processing. XXVII, 974 pages. 2007.

Vol. 4639: E. Csuhaj-Varjú, Z. Ésik (Eds.), Fundamentals of Computation Theory. XIV, 508 pages. 2007.

Vol. 4638: T. Stützle, M. Birattari, H. H. Hoos (Eds.), Engineering Stochastic Local Search Algorithms. X, 223 pages. 2007.

Vol. 4628: L.N. de Castro, F.J. Von Zuben, H. Knidel (Eds.), Artificial Immune Systems. XII, 438 pages. 2007.

Vol. 4627: M. Charikar, K. Jansen, O. Reingold, J.D.P. Rolim (Eds.), Approximation, Randomization, and Combinatorial Optimization. XII, 626 pages. 2007.

Vol. 4624: T. Mossakowski, U. Montanari, M. Haveraaen (Eds.), Algebra and Coalgebra in Computer Science. XI, 463 pages. 2007.

Vol. 4619: F. Dehne, J.-R. Sack, N. Zeh (Eds.), Algorithms and Data Structures. XVI, 662 pages. 2007.

Vol. 4618: S.G. Akl, C.S. Calude, M.J. Dinneen, G. Rozenberg, H.T. Wareham (Eds.), Unconventional Computation. X, 243 pages. 2007.

Vol. 4616: A. Dress, Y. Xu, B. Zhu (Eds.), Combinatorial Optimization and Applications. XI, 390 pages. 2007.

Vol. 4613: F.P. Preparata, Q. Fang (Eds.), Frontiers in Algorithmics. XI, 348 pages. 2007.

Vol. 4600: H. Comon-Lundh, C. Kirchner, H. Kirchner (Eds.), Rewriting, Computation and Proof. XVI, 273 pages. 2007.

Vol. 4599: S. Vassiliadis, M. Berekovic, T.D. Hämäläinen (Eds.), Embedded Computer Systems: Architectures, Modeling, and Simulation. XVIII, 466 pages. 2007.

Vol. 4598: G. Lin (Ed.), Computing and Combinatorics. XII, 570 pages. 2007.

Vol. 4596: L. Arge, C. Cachin, T. Jurdziński, A. Tarlecki (Eds.), Automata, Languages and Programming. XVII, 953 pages. 2007.

Vol. 4595: D. Bošnački, S. Edelkamp (Eds.), Model Checking Software. X, 285 pages. 2007.

Vol. 4590: W. Damm, H. Hermanns (Eds.), Computer Aided Verification. XV, 562 pages. 2007.

Vol. 4588: T. Harju, J. Karhumäki, A. Lepistö (Eds.), Developments in Language Theory. XI, 423 pages. 2007.

Vol. 4583: S.R. Della Rocca (Ed.), Typed Lambda Calculi and Applications. X, 397 pages. 2007.

Vol. 4580: B. Ma, K. Zhang (Eds.), Combinatorial Pattern Matching. XII, 366 pages. 2007.

Vol. 4576: D. Leivant, R. de Queiroz (Eds.), Logic, Language, Information and Computation. X, 363 pages. 2007.

Vol. 4547: C. Carlet, B. Sunar (Eds.), Arithmetic of Finite Fields. XI, 355 pages. 2007.

Vol. 4546: J. Kleijn, A. Yakovlev (Eds.), Petri Nets and Other Models of Concurrency – ICATPN 2007. XI, 515 pages. 2007.

Vol. 4545: H. Anai, K. Horimoto, T. Kutsia (Eds.), Algebraic Biology. XIII, 379 pages. 2007.

Vol. 4533: F. Baader (Ed.), Term Rewriting and Applications. XII, 419 pages. 2007.

Vol. 4528: J. Mira, J.R. Álvarez (Eds.), Nature Inspired Problem-Solving Methods in Knowledge Engineering, Part II. XXII, 650 pages. 2007.

Vol. 4527: J. Mira, J.R. Álvarez (Eds.), Bio-inspired Modeling of Cognitive Tasks, Part I. XXII, 630 pages. 2007.

Vol. 4525: C. Demetrescu (Ed.), Experimental Algorithms. XIII, 448 pages. 2007.

Vol. 4514: S.N. Artemov, A. Nerode (Eds.), Logical Foundations of Computer Science. XI, 513 pages. 2007.

Vol. 4513: M. Fischetti, D.P. Williamson (Eds.), Integer Programming and Combinatorial Optimization. IX, 500 pages. 2007.

Vol. 4510: P. Van Hentenryck, L.A. Wolsey (Eds.), Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems. X, 391 pages. 2007.

Vol. 4507: F. Sandoval, A.G. Prieto, J. Cabestany, M. Graña (Eds.), Computational and Ambient Intelligence. XXVI, 1167 pages. 2007.

Vol. 4501: J. Marques-Silva, K.A. Sakallah (Eds.), Theory and Applications of Satisfiability Testing – SAT 2007. XI, 384 pages. 2007.

Vol. 4497: S.B. Cooper, B. Löwe, A. Sorbi (Eds.), Computation and Logic in the Real World. XVIII, 826 pages. 2007.

Vol. 4494: H. Jin, O.F. Rana, Y. Pan, V.K. Prasanna (Eds.), Algorithms and Architectures for Parallel Processing. XIV, 508 pages. 2007.

Vol. 4493: D. Liu, S. Fei, Z. Hou, H. Zhang, C. Sun (Eds.), Advances in Neural Networks – ISNN 2007, Part III. XXVI, 1215 pages. 2007.

Vol. 4492: D. Liu, S. Fei, Z. Hou, H. Zhang, C. Sun (Eds.), Advances in Neural Networks – ISNN 2007, Part II. XXVII, 1321 pages. 2007.

Vol. 4491: D. Liu, S. Fei, Z.-G. Hou, H. Zhang, C. Sun (Eds.), Advances in Neural Networks – ISNN 2007, Part I. LIV, 1365 pages. 2007.

Vol. 4490: Y. Shi, G.D. van Albada, J.J. Dongarra, P.M.A. Sloot (Eds.), Computational Science – ICCS 2007, Part IV. XXXVII, 1211 pages. 2007.

Vol. 4489: Y. Shi, G.D. van Albada, J.J. Dongarra, P.M.A. Sloot (Eds.), Computational Science – ICCS 2007, Part III. XXXVII, 1257 pages. 2007.

Vol. 4488: Y. Shi, G.D. van Albada, J.J. Dongarra, P.M.A. Sloot (Eds.), Computational Science – ICCS 2007, Part II. XXXV, 1251 pages. 2007.

Vol. 4487: Y. Shi, G.D. van Albada, J.J. Dongarra, P.M.A. Sloot (Eds.), Computational Science – ICCS 2007, Part I. LXXXI, 1275 pages. 2007.

Vol. 4484: J.-Y. Cai, S.B. Cooper, H. Zhu (Eds.), Theory and Applications of Models of Computation. XIII, 772 pages. 2007.

Vol. 4475: P. Crescenzi, G. Prencipe, G. Pucci (Eds.), Fun with Algorithms. X, 273 pages. 2007.

Vol. 4474: G. Prencipe, S. Zaks (Eds.), Structural Information and Communication Complexity. XI, 342 pages. 2007.

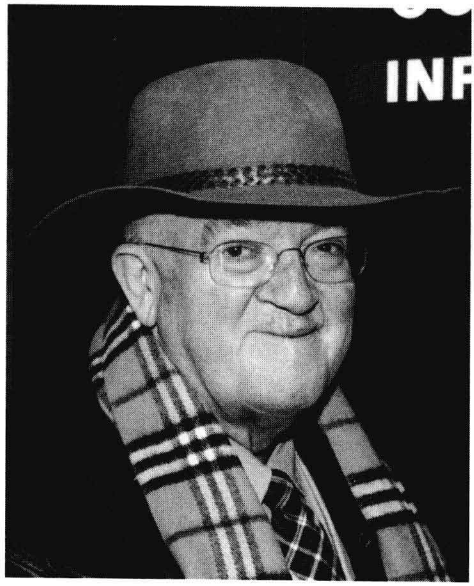
Vol. 4459: C. Cérin, K.-C. Li (Eds.), Advances in Grid and Pervasive Computing. XVI, 759 pages. 2007.

Vol. 4449: Z. Horváth, V. Zsók, A. Butterfield (Eds.), Implementation and Application of Functional Languages. X, 271 pages. 2007.

Vol. 4448: M. Giacobini (Ed.), Applications of Evolutionary Computing. XXIII, 755 pages. 2007.

Vol. 4447: E. Marchiori, J.H. Moore, J.C. Rajapakse (Eds.), Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics. XI, 302 pages. 2007.

Vol. 4446: C. Cotta, J.I. van Hemert (Eds.), Evolutionary Computation in Combinatorial Optimization. XII, 241 pages. 2007.



Dines Bjørner

Zhou Chaochen

Foreword

Two outstanding computer scientists will soon reach their 70th birthdays: Dines Bjørner was born on October 4, 1937 in Denmark and Zhou Chaochen was born on November 1, in the same year in China. To celebrate their birthdays, we present three LNCS volumes in their honour.

- *Formal Methods and Hybrid Real-Time Systems. Essays in Honour of Dines Bjørner and Zhou Chaochen on the Occasion of Their 70th Birthdays.* Papers presented at a Symposium held in Macao, China, September 24–25, 2007. LNCS volume **4700**. Springer 2007.
- *Domain Modelling and the Duration Calculus.* International Training School, Shanghai, China, September 10–21, 2007. Advanced Lectures. LNCS volume 4710. Springer 2007.
- *Theoretical Aspects of Computing* - ICTAC 2007. 4th International Colloquium, Macao, China, September 26–28, 2007, Proceedings. LNCS volume 4711. Springer 2007.

DINES BJØRNER is known for his many contributions to the theory and practice of formal methods for software engineering. He is particularly associated with two formal methods, although his influence is far wider. He worked with Cliff Jones and others on the *Vienna Development Method (VDM)*, initially at IBM in Vienna. Later, he was involved in producing the *Rigorous Approach to Industrial Software Engineering (RAISE)* formal method with tool support. His three-volume *magnum opus* on software engineering covers *Abstraction and Modelling*, *Specification of Systems and Languages*, and *Domains, Requirements, and Software Design*. He was a professor at the Technical University of Denmark (DTU) in Lyngby, near Copenhagen. He was the founding director of the United Nations University International Institute for Software Technology (UNU-IIST) in Macao during the 1990s. He was a co-founder of VDM-Europe, which transformed to become Formal Methods Europe, an organisation that promotes the use of formal methods. Its 18 monthly symposia have become the leading academic events in formal methods. Dines Bjørner is a Knight of the Order of the Dannebrog and was awarded the John von Neumann Medal in Budapest in 1994. He received a Doctorate (*honoris causa*) from the Masaryk University in Brno in 2004. He is a Fellow of both the IEEE and the ACM.

ZHOU CHAOCHEN is known for his seminal contributions to the theory and practice of timed and hybrid systems. His distinguished academic career started as an undergraduate in mathematics and mechanics at Peking University (1954–58) and as a postgraduate at the Institute for Computing Technology in the Chinese Academy of Sciences (1963–67). He continued his career at Peking University and the Chinese Academy, until he made an extended visit to Oxford University

Computing Laboratory (1989–92) at the invitation of Sir Tony Hoare FRS. Here he was the prime instigator of *Duration Calculus*, an interval logic for real-time systems, developed as part of a European ESPRIT project on Provably Correct Systems. He made further extended visits during the periods 1990–92 and 1995–96, as a visiting professor at the Technical University of Denmark, Lyngby, at the invitation of Dines Bjørner. He was a Principal Research Fellow at UNU-IIST during the period 1992–97, before becoming its director, an appointment he held from 1997 to 2002. He is a member of the Chinese Academy of Sciences and the Third World Academy of Sciences.

We thank both Dines Bjørner and Zhou Chaochen for their years of generous, wise advice, to us and to their many other colleagues, students, and friends. They have both been unfailingly inspiring, enthusiastic, and encouraging.

July 2007 J.C.P.W.

Tabula Gratulatoria

Nazareno Aguirre
Bogdan Aman
Damian Barsotti
Marc Bezem
Nikolaj Bjørner
Javier Blanco
Guillaume Bonfante
Pontus Boström
Aske Wiid Brekling
Jan Bretschneider
Alan Burns
Andrew Butterfield
Zining Cao
Pablo Castro
Haiyan Chen
Huowang Chen
Yinghua Chen
Zhenbang Chen
Gabriel Ciobanu
Robert Colvin
Pascal Coupey
Werner Damm
Dang Van Hung
Rafael del Vado Vírseda
Fredrik Degerlund
Catalin Dima
Wei Dong
Brijesh Dongol
Asger Eir
Estevez Elsa
Ignacio Fábregas
Dirk Fahland
John Fisher
John Fitzgerald
Christophe Fouqueré
Leo Freitas
David Frutos Escrig
Kokichi Futatsugi
Chris George
Michael Reichhardt Hansen
Anne Haxthausen
Ian Hayes
He Jifeng
Michael A. Jackson
Tomasz Janowski
Cliff Jones
Mathai Joseph
Takashi Kitamura
John Knudsen
Maciej Koutny
Padmanabhan Krishnan
Hans Langmaack
Ruggero Lanotte
Alessandro Lapadula
Peter Gorm Larsen
Martin Leucker
Jing Li
Li Xiaoshan
Huimin Lin
Xiang Ling
Daguang Liu
Wanwei Liu
Xinxin Liu
Zhiming Liu
Jean-Vincent Loddo
Niels Lohmann
Roussanka Loukanova
Jan Madsen
Tom Maibaum
Dino Mandrioli
Jean-Yves Marion
Peter Massuthe
Andrea Matta
Alfred Mikschl
Lionel Morel
Peter Mosses
Masaki Nakamura
Virginia Niculescu
Thomas Noll
Jens Oehlerking
Ernst-Rüdiger Olderog
Romain Péchoux
Miguel Palomino
Jun Pang
Jan Peleska
Martin Pěnička
André Platzer
Rosario Pugliese
Brian Randell
Silvio Ranise
Anders Ravn
Wolfgang Reisig
Stefan Rieger
Matteo Rossi
Cesar Sanchez
J.W. Sanders
Christelle Scharff
Marc Segelken
Quirico Semeraro
Kaisa Sere
Arne Skou
Paola Spoletini
Christian Stahl
Volker Stolz
K. Subramani
Francesco Tiezzi
Tullio Tolio
Marina Waldén
Ji Wang
Boris Wirtz
Jim Woodcock
Peng Wu
Zhilin Wu
Bican Xia
Lu Yang
Lu Yang
Naijun Zhan
Huibiao Zhu

Preface

This volume contains the papers presented at the *Festschrift Symposium* held September 24–25, 2007 in Macao on the occasion of the 70th birthdays of Dines Bjørner and Zhou Chaochen. It consists of 25 papers written by 59 authors. Online conference management was provided by EasyChair.

It is now difficult to remember exactly when it came to us that we should organise a celebration for the 70th birthdays of Dines Bjørner and Zhou Chaochen, which happily coincide this year. But I do know that the idea was a popular one. Zhiming Liu suggested that we should organise the symposium as part of the International Colloquium on Theoretical Aspects of Computing, which seemed perfect given that this series was founded by UNU/IIST. The event quickly took shape as He Jifeng offered to host a Training School in Shanghai with the assistance of Chris George, Geguang Pu, and Yong Zhou, and Cliff Jones agreed to help with the academic organisation of the symposium and the colloquium. Everything then just fell into place, thanks to the excellent help provided by the local organisers in Macao and Shanghai.

The subjects for the lectures for the school were obvious to us all: two topics pioneered by Dines Bjørner and Zhou Chaochen, both currently very active research areas. For the *Festschrift Symposium*, authors were invited to write on an original topic of their choosing. And for the colloquium, a general call-for-papers resulted in a satisfying collection of rigorously reviewed papers in theoretical computer science, including automata theory, case studies, concurrency, real-time systems, semantics and logics, and specification and verification.

So we have ended up with three volumes, one each for the school, symposium, and colloquium, which collectively amount to some 1,300 pages. And still there was not enough room for the many additional distinguished names we would have liked to invite.

To Dines and Chaochen from all of us:

We hope that you enjoy reading these books.

Happy birthday to both of you!

June 2007 J.C.P.W.

Organization

Programme Chairs

Cliff Jones
Zhiming Liu
Jim Woodcock

Local Organization

Kitty Chan
Wendy Hoi
Chris George
Violet Pun

Table of Contents

Models and Software Model Checking of a Distributed File Replication System

Nikolaj Bjørner

Microsoft Research, One Microsoft Way, Redmond, WA, 98074, USA
nbjorner@microsoft.com

Abstract. With the Distributed File System Replication component, DFS-R, as the central theme, we present selected protocol problems and validation methods encountered during design and development. DFS-R is currently deployed in various contexts; in Windows Server 2003-R2, Windows Live Messenger (Sharing Folders), and Windows Vista (Meeting spaces). The journey from an initial design sketch to a shipped product required mainly the dedicated effort of several testers, developers, program managers, and several others; but in some places cute problems related to distributed consensus and software model-checking emerged. This paper presents a few of these, including a distributed garbage collection problem, distributed consensus problems for reconciling tree-like data structures, using model-based test case generation, and the use of software model checking in design and development process.

1 Introduction

Designing and building distributed systems is challenging, especially if they need to scale, perform, satisfy customer functionality requirements, and, oh well, work. An example of a particularly challenging distributed system is multi-master, optimistic, file replication. One of the distinguished factors making distributed file replication hard is that file replication comes with a very substantial data component: the protocols need to be sufficiently aware of file system semantics, such as detecting and resolving name conflicting file creates and concurrent updates. Such races are just the tip of the iceberg. In comparison, cache coherence protocols that are known to be challenging to design, have a trivial data component, but to be fair have stronger consistency requirements.

Subtle protocol bugs can go (and have indeed gone) undetected for years due to the large number of interactions that are possible. With a sufficient number of deployments they *will* be encountered in the field, have costly consequences, and be extremely challenging to analyze. Our experience in developing DFS-R from the bottom up, is used to demonstrate several complementary uses of model-based techniques for system design and exploration. This paper provides an experience report on these selected methods. Note that the material presented here reflect only a very partial view of the design and test of DFS-R.

DFS-R was developed to address correctness, scale, and management challenges encountered with a predecessor file replication product. Thus, the original

C.B. Jones, Z. Liu, J. Woodcock (Eds.): Bjørner/Zhou Festschrift, LNCS 4700, pp. 1–23, 2007.

impression was that we had the luxury of tackling a relatively well defined problem; to build a replication system specifically handling features of the file system NTFS, for replicating files between globally dispersed branch offices of corporations. Later on, it would turn out that DFS-R could be embedded within other scenarios, such as, in an instant messenger product. However, we consciously avoided over-loading with features from the onset. It means that DFS-R, for instance does not replicate files synchronously, only asynchronously (as it is meant for wide area networks); does not replicate general directed acyclic graphs, only tree-like structure; and does not maintain fine-grained tracking of operations, only state. While several such problems are interesting in other contexts, they did not fall into the scope of our original goals.

The organization of this paper follows the top-down design flow of DFS-R. The DFS-R system was originally conceived as a strictly state-based file replication protocol. Section 2 elaborates on the differences between state-based and operations-based replication systems. We developed a high-level state machine specification of DFS-R by using a transition system presented as a collection of guarded commands. The guarded commands were subsequently implemented as an applicative program in OCaml. This paved the way for performing efficient state space exploration on top of the design. Section 3 elaborates on the protocol, and Section 4 summarizes prototyping experiences. As the development took place, several assumptions made in the abstract design turned out to be unrealistic, and we redid the high-level design using the AsmL tools that were built at Microsoft for software modeling and test case generation. Section 5 elaborates on the experiences from using AsmL. A number of well-separated distributed protocol problems emerged during the development. Section 6 describes the distributed tree reconciliation problem, and how we used a model checker, Zing, to expose bugs in both protocol proposals and legacy implementations. Section 7 describes the distributed tombstone garbage collection problem and a solution to it. While one cannot expect to get anywhere without a high-level understanding of the protocols involved in DFS-R, it is equally unrealistic to expect developing a production quality system without addressing systems problems. We were thus faced with a potentially large gap between simplified protocol substrates and the production code. Encouraged by the ability of the model-based state space exploration to expose subtle interaction bugs we repeated the state space exploration experiment on top of the production core. The resulting backtracking search tool may best be characterized as a hybrid software model checking, run-time verification tool. It operates directly at the source code level. It uses techniques, such as partial order reduction to prune search and custom allocation routines to enable backtracking search. Section 8 describes the infrastructure we developed and the experiments covering $\frac{1}{2}$ trillion scenarios.

2 File Replication

The style of replication systems under which DFS-R falls into is surveyed extensively in [1]. We here summarize a few of the main concepts relevant for

DFS-R. The problem that DFS-R solves is to maintain mirror copies of selected directories across large networks of servers. The directories that are selected for replication are called *replicated folders.* Files and directories within these directories may be created, modified, deleted, moved, or renamed at any of the mirror sites. It is the job of DFS-R to distribute changes, detect and reconcile conflicts automatically when they arise. Distributed replication systems can be categorized according to what problems they solve and how they solve them. Figure 1 summarizes some of the main design choices one has when designing a replication system.

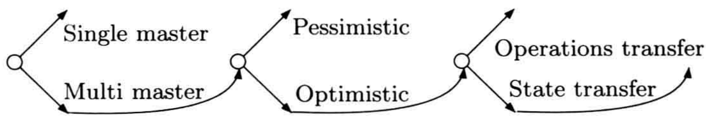


Fig. 1. Replication system ontologies

Multi Master Replication. DFS-R is a multi-master replication system. Any machine may participate in changing resources, and their updates will have to be reconciled with updates from any other machine. A (selective) single-master system only replicates changes from a set of selected machines. All other machines are expected to maintain a mirror copy of the masters. This would mean that file system changes on non-masters would have to be reverted. If there is a designated master, one can even choose to maintain truth centrally. The challenge there is managing fail-over and network disconnects.

Optimistic Replication. To support wide area networks (spanning the globe) DFS-R supports optimistic updates to files. This means that any machine may submit updates to resources without checking first whether the update is in conflict with other updates. Pessimistic replication schemes avoid concurrent update conflicts by serializing read and write operations using locking schemes.

State and Operation Transfer. A file system state is the result of the file operations (create, update, delete, move) that are performed on it. This suggests two approaches to realize file replication: intercept and replay the file operations, called operation transfer, or capture the file system state and replicate it as it is, called state transfer. DFS-R implements a state transfer protocol. There are several hard challenges with operations-transfer based systems. One is merging operations into a consistent serialization. Another, is space, as operations are not necessarily amenable to garbage collection.

Perspective. There is no single choice of design parameters that handles all customer scenarios. In some configurations, corporations wish to designate machines as read-only, and can manage the additional constraints this leaves on