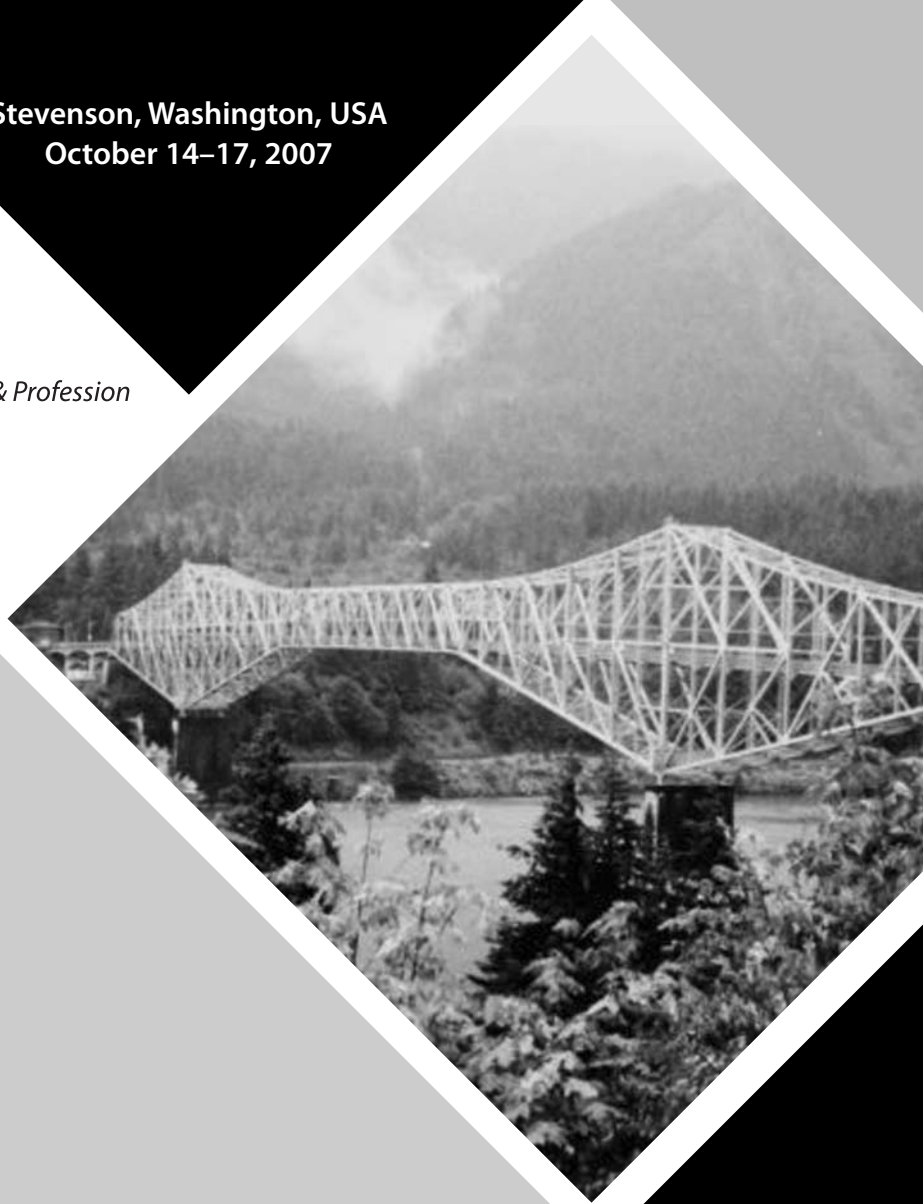


Stevenson, Washington, USA
October 14–17, 2007



Association for
Computing Machinery

Advancing Computing as a Science & Profession



SOSP'07

Proceedings of the 21st ACM Symposium on
Operating Systems Principles

Sponsored by:

ACM SIGOPS



**Association for
Computing Machinery**

Advancing Computing as a Science & Profession

**The Association for Computing Machinery
2 Penn Plaza, Suite 701
New York, New York 10121-0701**

Copyright © 2007 by the Association for Computing Machinery, Inc. (ACM). Permission to make digital or hard copies of portions of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyright for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permission to republish from: Publications Dept., ACM, Inc. Fax +1 (212) 869-0481 or <permissions@acm.org>.

For other copying of articles that carry a code at the bottom of the first or last page, copying is permitted provided that the per-copy fee indicated in the code is paid through the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

Notice to Past Authors of ACM-Published Articles

ACM intends to create a complete electronic archive of all articles and/or other material previously published by ACM. If you have written a work that has been previously published by ACM in any journal or conference proceedings prior to 1978, or any SIG Newsletter at any time, and you do NOT want this work to appear in the ACM Digital Library, please inform permissions@acm.org, stating the title of the work, the author(s), and where and when published.

ISBN: 978-1-59593-591-5

Additional copies may be ordered prepaid from:

ACM Order Department
PO Box 11405
New York, NY 10286-1405

Phone: 1-800-342-6626
(US and Canada)
+1-212-626-0500
(all other countries)
Fax: +1-212-944-1318
E-mail: acmhelp@acm.org

ACM Order Number 83300706

Printed in the USA

Foreword

Please enjoy the proceedings of the 21st ACM Symposium on Operating Systems Principles–SOSP’07. In the SOSP tradition, the 25 papers herein explore a wide range of computer systems topics, including traditional ones such as concurrency as well as new ones such as “hardening” Web browsers. Collectively these papers report on some of the most creative and thought-provoking ideas in computer systems today and how they work out in practice. The 25 papers were shepherded by PC members to ensure that they are easy to read. We hope you will enjoy learning from these papers.

Selecting 25 papers out of 131 submissions was difficult because so many of the submissions were of high quality. To make the selection process as fair and as consistent as possible the program committee employed a different process than used by previous SOSPs (but used successfully by other conferences such as SIGCOMM). The program committee consisted of 13 “heavy”-load and 13 “light”-load members. The heavy-load members reviewed about 34 submissions each and attended the face-to-face PC meeting in Cambridge, MA USA. The light-load members reviewed about 24 papers each and did not attend the PC meeting. In contrast, recent SOSPs used a small number of PC members (12-15) who read a large fraction of all submissions, sometimes assisted by external reviewers. SOSPs before that required all PC members to read all submissions.

The goal of the new process was to resolve the tension between having high-quality, consistent reviews, a large number of submissions (it has been steadily growing over the years), and a productive face-to-face meeting. With more PC members the PC did not have to rely on external reviews, which can be inconsistent because the external reviewers see only a small sample of the submissions, yet the workload for the individual PC members was manageable, allowing thorough reviewing. By having a subset of the PC members meet in person, the PC was able to have in-depth discussion and reach consensus through discussion (rather than voting). The larger overall PC also allowed a broader group of people to participate in the decisions.

Paper selection was a three round process, with multiple reviews by the PC generated in each round and with reviewers targeted by subject expertise. The first two rounds reduced the pool of considered papers by 50%. The 62 remaining papers produced another two reviews apiece and all 705 reviews were assessed in preparation for the PC meeting. At the PC meeting, the 62 papers were ranked by review scores for discussion order and each assigned a champion to summarize content and strengths and to lead the discussion on individual papers. The PC discussion for each paper followed until consensus was reached. Throughout the process anonymity was maintained and conflicts of interest precluded by removing authors or those with direct association with an author from the discussion. In the final selection, 3 papers were co-authored by heavy-load PC members, and 6 were co-authored by light-load PC members.

Did the PC make good decisions? This question is probably best answered by you after reading the papers! It is interesting to note, however, that a shadow PC chaired and organized by Rebecca Isaacs (Microsoft Research, Cambridge, England) reviewed 101 of the 131 submissions (which included 18 of the 25 papers accepted by the real PC) and accepted 16 papers. Of the 18 papers accepted by the real PC, 9 were accepted by the shadow PC, 4 were discussed by the shadow PC, and 5 didn’t make it to the discussion at the shadow PC meeting (the shadow PC discussed 40 submissions). An informal review suggests that the variations in decisions were partially due to the fact that the shadow PC’s goals were different from the real PC’s. The shadow PC’s main goal was to educate participants about how a PC works, how to review papers, etc. and members volunteered to participate; the real PC members were carefully chosen to provide both depth and breadth across a

wide range of topics. This difference in focus resulted in a few important modifications to the decision process: the shadow PC members produced 4 reviews per submission and saw fewer submissions, had less time to absorb the reviews before the meeting, and had less expertise in certain areas. A full report will be submitted to SIGOPS Operating Systems Review.

A successful conference goes beyond the accepted papers, building and supporting its community. At SOSP this year, and in celebration of the 20th anniversary of the SYSTERS group, we have introduced two special programs. First, we recognize the importance of increasing the participation of women and underrepresented minorities in systems research. And to be successful, this participation has to reach to undergraduates and show them the excitement and interesting problems in systems. Toward this end, we established an additional scholarship opportunity, supported by industry contributors, that has supported these targeted groups to attend SOSP. Second, the participation by women works best when undergraduates and graduates are shown the way by women already participating in the field. For this, we created a special one-day workshop for women to develop this community as a prelude to the beginning of the SOSP conference. Support from our industry contributors and from NSF and CRA-W has been outstanding. Equally impressive has been the support from the organizing team and all who have made this possible. These initiatives have been embraced enthusiastically.

SOSP is a great conference mostly because it attracts so many high-quality submissions, and we would like to thank all the authors who submitted. We would also like to thank the PC members for the tremendous amount of work they did: reviewing the submissions, providing constructive feedback, and shepherding the accepted submissions. Organizing a conference is a team effort, and we would like to thank the team: Jon Walpole (Local arrangements), Jay Lorch (Industry contributions), Jason Flinn (Publicity), Robbert van Renesse and Hakim Weatherspoon (Scholarships), Mike Kozuch (Registration), Eddie Kohler (HotCRP), Mema Roussopoulos (Posters), David Mazières (WIPs), and Carla Ellis, Sharon Perl, and Barbara Liskov (Women's workshop). All of these people have shown great dedication and have worked very hard to make this conference a success. Finally we would like to thank the sponsors—without their financial support the conference could not have happened.

We hope that you will find the SOSP program interesting and that the symposium will provide you with a valuable opportunity to share ideas with other researchers and practitioners from institutions around the world.

Tom Bressoud

*General Chair
Denison University*

Frans Kaashoek

*Program Chair
MIT*

Table of Contents

21st ACM Symposium on Operating Systems Principles Organization	vii
--	-----

Session 1: Web Meets Operating Systems

• Protection and Communication Abstractions for Web Browsers in MashupOS	1
Helen J. Wang, Xiaofeng Fan, Jon Howell (<i>Microsoft Research</i>), Collin Jackson (<i>Stanford University</i>)	
• AjaxScope: A Platform for Remotely Monitoring the Client-Side Behavior of Web 2.0 Applications	17
Emre Kiciman, Benjamin Livshits (<i>Microsoft Research</i>)	
• Secure Web Applications via Automatic Partitioning	31
Stephen Chong, Jed Liu, Andrew C. Myers, Xin Qi, K. Vikram, Lantian Zheng, Xin Zheng (<i>Cornell University</i>)	

Session 2: Byzantine Fault Tolerance

• Zyzyva: Speculative Byzantine Fault Tolerance	45
Ramakrishna Kotla, Lorenzo Alvisi, Mike Dahlin, Allen Clement, Edmund Wong (<i>University of Texas at Austin</i>)	
• Tolerating Byzantine Faults in Transaction Processing Systems using Commit Barrier Scheduling	59
Ben Vandiver, Hari Balakrishnan, Barbara Liskov, Sam Madden (<i>Massachusetts Institute of Technology</i>)	
• Low-Overhead Byzantine Fault-Tolerant Storage	73
James Hendricks, Gregory R. Ganger (<i>Carnegie Mellon University</i>), Michael K. Reiter (<i>University of North Carolina at Chapel Hill</i>)	

Session 3: Concurrency

• TxLinux: Using and Managing Hardware Transactional Memory in an Operating System	87
Christopher J. Rossbach, Owen S. Hofmann, Donald E. Porter, Hany E. Ramadan, Aditya Bhandari, Emmett Witchel (<i>University of Texas at Austin</i>)	
• MUVI: Automatically Inferring Multi-Variable Access Correlations and Detecting Related Semantic and Concurrency Bugs	103
Shan Lu, Soyeon Park, Chongfeng Hu, Xiao Ma, Weihang Jiang (<i>University of Illinois</i>), Zhenmin Li (<i>University of Illinois & CleanMake Inc.</i>), Raluca A. Popa (<i>Massachusetts Institute of Technology</i>), Yuanyuan Zhou (<i>University of Illinois & CleanMake Inc.</i>)	

Session 4: Software Robustness

• Bouncer: Securing Software by Blocking Bad Input	117
Manuel Costa, Miguel Castro, Lidong Zhou, Lintao Zhang (<i>Microsoft Research</i>), Marcus Peinado (<i>Microsoft</i>)	
• Triage: Diagnosing Production Run Failures at the User's Site	131
Joseph Tucek, Shan Lu, Chengdu Huang, Spiros Xanthos, Yuanyuan Zhou (<i>University of Illinois at Urbana Champaign</i>)	
• *iComment: Bugs or Bad Comments? *	145
Lin Tan, Ding Yuan, Gopal Krishna (<i>University of Illinois at Urbana Champaign</i>), Yuanyuan Zhou (<i>University of Illinois at Urbana Champaign & CleanMake Co.</i>)	

Session 5: Distributed Systems

• Sinfonia: A New Paradigm for Building Scalable Distributed Systems	159
Marcos K. Aguilera, Arif Merchant, Mehul Shah, Alistair Veitch (<i>Hewlett Packard Laboratories</i>), Christos Karamanokis (<i>VMware</i>)	
• PeerReview: Practical Accountability for Distributed Systems	175
Andreas Haeberlen (<i>Max Planck Institute for Software Systems and Rice University</i>), Petr Kouznetsov, Peter Druschel (<i>Max Planck Institute for Software Systems</i>)	

- **Attested Append-Only Memory: Making Adversaries Stick to their Word** 189
Byung-Gon Chun (*University of California at Berkeley*), Petros Maniatis (*Intel Research Berkeley*),
Scott Shenker (*University of California at Berkeley & ICSI*),
John Kubiatowicz (*University of California at Berkeley*)
- **Dynamo: Amazon’s Highly Available Key-value Store** 205
Giuseppe De Candia, Deniz Hastorun, Madan Jampani, Gunavardhan Kakulapati, Avinash Lakshman ,
Alex Pilchin, Swaminathan Sivasubramanian, Peter Voshall, Werner Vogels (*Amazon.com*)

Session 6: System Maintenance

- **Staged Deployment in Mirage, an Integrated Software Upgrade Testing and Distribution System** 221
Olivier Cramer, Nikola Knežević, Dejan Kostić (*EPFL*),
Ricardo Bianchini (*Rutgers University*), Willy Zwaenepoel (*EPFL*)
- **AutoBash: Improving configuration management with operating system causality analysis** 237
Ya-Yunn Su, Mona Attariyan, Jason Flinn (*University of Michigan*)

Session 7: Energy

- **Integrating Concurrency Control and Energy Management in Device Drivers** 251
Kevin Klues (*Stanford University, Washington University & Technische Universität Berlin*),
Vlado Handziski (*Technische Universität Berlin*), Chenyang Lu (*Washington University*),
Adam Wolisz (*Technische Universität Berlin & University of California at Berkeley*),
David Culler (*Arch Rock Corporation & University of California at Berkeley*),
David Gay (*Intel Research*), Philip Levis (*Stanford University*)
- **VirtualPower: Coordinated Power Management in Virtualized Enterprise Systems** 265
Ripal Nathuji, Karsten Schwan (*Georgia Institute of Technology*)

Session 8: Storage

- **DejaView: A Personal Virtual Computer Recorder** 279
Oren Laadan, Ricardo A. Baratto, Dan B. Phung, Shaya Potter, Jason Nieh (*Columbia University*)
- **Improving File System Reliability with I/O Shepherding** 293
Haryadi S. Gunawi (*University of Wisconsin*), Vijayan Prabhakaran (*Microsoft Research*),
Swetha Krishnan, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau (*University of Wisconsin*)
- **Generalized File System Dependencies** 307
Christopher Frost, Mike Mammarella, Eddie Kohler (*University of California at Los Angeles*),
Andrew de los Reyes (*Google*), Shant Hovsepian (*University of California at Los Angeles*),
Andrew Matsuoka (*University of Texas at Austin*), Lei Zhang (*Google*)

Session 9 Operating System Security

- **Information Flow Control for Standard OS Abstractions** 321
Maxwell Krohn, Alexander Yip, Micah Brodsky, Natan Cliffer, M. Frans Kaashoek
(*Massachusetts Institute of Technology*), Eddie Kohler (*University of California at Los Angeles*),
Robert Morris (*Massachusetts Institute of Technology*)
- **SecVisor: A Tiny Hypervisor to Provide Lifetime Kernel Code Integrity for Commodity OSes** 335
Arvind Seshadri, Mark Luk, Ning Qu, Adrian Perrig (*Carnegie Mellon University*)
- **Secure Virtual Architecture: A Safe Execution Environment for Commodity Operating Systems** 351
John Criswell, Andrew Lenharth (*University of Illinois at Urbana-Champaign*),
Dinakar Dhurjati (*DoCoMo Communications Laboratories*),
Vikram Adve (*University of Illinois at Urbana Champaign*)

- **Author Index** 367

21st ACM Symposium on Operating Systems Principles

Organization

General Chair & Treasurer: Thomas C. Bressoud (*Denison University, USA*)

Program Chair: M. Frans Kaashoek (*Massachusetts Institute of Technology, USA*)

Financial Support Chair: Jacob R. Lorch (*Microsoft Research, USA*)

Local Arrangements Chair: Jonathan Walpole (*Portland State University, USA*)

Registration Chair: Michael Kozuch (*Intel Research, USA*)

Publicity Chair: Jason Flinn (*University of Michigan, USA*)

SOSP Women's Workshop: Carla Ellis (*Duke University, USA*)
Sharon Perl (*Google, USA*)
Barbara Liskov (*Massachusetts Institute of Technology, USA*)

Student Scholarship Committee: Robbert van Renesse (*Cornell University, USA*)
Hakim Weatherspoon (*Cornell University, USA*)

Poster Session Chair: Mema Roussopoulos (*Harvard University, USA*)

Work-in-progress Session Chair: David Mazières (*Stanford University, USA*)

HotCRP Czar: Eddie Kohler (*University of California Los Angeles, USA*)

Shadow PC Chair: Rebecca Isaacs (*Microsoft Research, England*)

Hall of Fame Award Committee: Ken Birman (*Cornell University, USA*)
Peter Chen (*University of Michigan, USA*)
Peter Druschel (*MPI-SWS, Germany*)
Mike Jones (*Microsoft Research, USA*)
Frans Kaashoek (*Massachusetts Institute of Technology, USA*)
Jeff Mogul (*Chair, HP Labs, USA*)
Larry Peterson (*Princeton University, USA*)
John Wilkes (*HP Labs, USA*)

Program Committee: Andrea Arpaci-Dusseau (*University of Wisconsin, USA*)
Paul Barham (*Microsoft Research, England*)
Peter Chen (*University of Michigan, USA*)
Jeff Dean (*Google, USA*)
Richard Draves (*Microsoft Research, USA*)
Peter Druschel (*Max Planck Institute for Software Systems, Germany*)
Carla Ellis (*Duke University, USA*)
Hermann Härtig (*Dresden University of Technology, Germany*)
M. Frans Kaashoek (*Massachusetts Institute of Technology, USA*)
Orran Krieger (*IBM, USA*)
Hank Levy (*University of Washington, USA*)
David Mazières (*Stanford University, USA*)
Andrew Myers (*Cornell University, USA*)
Jeff Mogul (*HP Labs, USA*)
Vivek Pai (*Princeton University, USA*)
Adrian Perrig (*Carnegie Mellon University, USA*)
Sharon Perl (*Google, USA*)
Sylvia Ratnasamy (*Intel Research, USA*)
Mendel Rosenblum (*Stanford University, USA*)
Mema Roussopoulos (*Harvard University, USA*)
Stefan Savage (*University of California San Diego, USA*)
Michael Schroeder (*Microsoft Research, USA*)
Ion Stoica (*University of California Berkeley, USA*)
Yuanyuan Zhou (*University of Illinois, USA*)
WeiMin Zheng (*Tsinghua University, China*)

Additional Reviewers:

Krste Asanovic	Alan Mislove
Hari Balakrishnan	Robert Morris
Cullen Bash	Robert O’Callahan
Lujo Bauer	Bryan Parno
Micah Brodsky	Ansley Post
David Brumley	Charles Reis
Tanya Bragin	Mike Reiter
Mike Burrows	Partha Ranganathan
Russ Cox	Andrey Rybalchenko
Rodrigo Fonseca	Arvind Seshadri
Bryan Ford	Atul Singh
Armando Fox	Emil Sit
Greg Ganger	Dawn Song
Steven Gribble	Jacob Strauss
Andreas Haeberlen	Jeremey Stribling
Max Krohn	K. Vikram
Chris Lesniewski-Laas	Benjamin Wester
Dave Levin	Alex Yip
Z. Morley Mao	Lantian Zheng
Petros Maniatis	Xin Zheng
Jon McCune	

Sponsor:  **SIGOPS**
ACM SIG on Operating Systems

Supporters: 
POWERED BY INTELLECT
DRIVEN BY VALUES



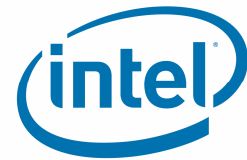
Microsoft



Google



 vmware



IBM

