

8th USENIX Symposium on Operating Systems Design and Implementation (OSDI'08)

San Diego, California, USA
8 - 10 December 2008

ISBN: 978-1-7138-0462-8

Printed from e-media with permission by:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571



Some format issues inherent in the e-media version may also appear in this print version.

Copyright© (2008) by Usenix Association
All rights reserved.

Printed with permission by Curran Associates, Inc. (2020)

For permission requests, please contact Usenix Association
at the address below.

Usenix Association
2560 Ninth Street, Suite 215
Berkeley, California, 94710

<https://www.usenix.org/>

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2633
Email: curran@proceedings.com
Web: www.proceedings.com

Symposium Organizers

Program Co-Chairs

Richard Draves, *Microsoft Research*
Robbert van Renesse, *Cornell University*

Program Committee

Marcos Aguilera, *Microsoft Research*
Lorenzo Alvisi, *University of Texas, Austin*
Remzi Arpaci-Dusseau, *University of Wisconsin, Madison*
Eric Brewer, *University of California, Berkeley, and Intel Research Berkeley*
Brad Chen, *Google, Inc.*
Fred Douglass, *IBM Research*
Greg Ganger, *Carnegie Mellon University*
Galen Hunt, *Microsoft Research*
Anthony Joseph, *University of California, Berkeley*
Dina Katabi, *MIT*
Kim Keeton, *HP Labs*
Idit Keidar, *Technion*
Terence Kelly, *HP Labs*
Dejan Kostić, *École Polytechnique Fédérale de Lausanne*
Philip Levis, *Stanford University*
David Lie, *University of Toronto*
Jack Lo, *VMware*
Dahlia Malkhi, *Microsoft Research*
Erich Nahum, *IBM Research*
Fernando Pedone, *University of Lugano*

Ian Pratt, *University of Cambridge*
Dave Presotto, *Google, Inc.*
Krithi Ramamritham, *IIT Bombay*
Rodrigo Rodrigues, *Max Planck Institute for Software Systems*
Wolfgang Schröder-Preikschat, *University of Erlangen-Nürnberg*
Marvin Theimer, *Amazon.com*
Leendert van Doorn, *AMD*
Geoffrey M. Voelker, *University of California, San Diego*
Jim Waldo, *Sun Microsystems, Inc.*
Helen Wang, *Microsoft Research*
Hakim Weatherspoon, *Cornell University*

Poster and Work-in-Progress (WiP) Sessions Co-Chairs

Dejan Kostić, *École Polytechnique Fédérale de Lausanne*
Philip Levis, *Stanford University*

Steering Committee

Brian Bershad, *University of Washington*
Jeff Mogul, *HP Labs*
Margo Seltzer, *Harvard University*
Ellie Young, *USENIX*

The USENIX Association Staff

External Reviewers

Atul Adya	Peter Druschel	Jim Larus	Alex Shraer
Sharad Agarwal	Dawson Engler	Charles R. Lefurgy	Craig Soules
Aditya Akella	Kave Eshghi	Harry Li	Chris Stewart
Martin Arlitt	Ittai Eyal	Mark Lillibridge	Marc Stiegler
Gal Badishi	Pedro Fonseca	Roie Melamed	Ram Swaminathan
Mahesh Balakrishnan	Alex Friedman	Alan Mislove	Chunqiang Tang
Paul Barham	Moises Goldsmidt	Tomer Morad	Doug Terry
Amit Berman	Albert Greenberg	Brad Morrey	Nedeljko Vasic
Mark Bickford	Chris Grier	Robert Morris	Luis Veiga
Sebastian Biemüller	Saikat Guha	Madanial Musuvathi	Michael Vrabie
Hans Boehm	Ramki Gummadi	Andrew Myers	Ollie Williams
Bill Bolosky	Diwaker Gupta	Pradeep Padala	Ted Wobber
Edward Bortnikov	Maxim Gurevich	Yoonho Park	Yinglian Xie
Dhruva Chakrabarti	Zvika Guz	Matt Piotrowski	Praveen Yalagandula
Anton Chernoff	Andreas Haeberlen	Ansley Post	Junfeng Yang
Stephen Chong	Gernot Heiser	Tom Rodenheffer	Jian Yin
Allen Clement	Maurice Herlihy	Andrey Rybalchenko	Ken Yocum
Olivier Crameri	Michael Hicks	Jim Saxe	Fang Yu
Uwe Dannowski	Michael Hohmuth	Nicolas Schiper	Ming Zhang
Stephan Diestelhorst	Aman Kansal	Simon Schubert	
John Douceur	Alan Karp	David Schultz	

For additional copies of these proceedings contact:

USENIX Association, 2560 Ninth Street, Suite 215, Berkeley, CA 94710 USA
Phone: 510-528-8649 • FAX: 510-548-5738 • office@usenix.org • http://www.usenix.org

The price is \$50 for members and \$60 for nonmembers.
Outside the U.S.A. and Canada, please add \$20 per copy for postage (via air printed matter).

Thanks to Our Sponsors

Platinum Sponsor



Reception and Bronze Sponsors



Gold Sponsor



Silver Sponsors



Conference Network Sponsor



Bronze Sponsors



Media Sponsors

ACM Queue
Addison-Wesley Professional/
Prentice Hall Professional/
Cisco Press

InfoSec News
ITtoolbox

Linux Journal

Linux Pro Magazine
LXer.com
StorageNetworking.org
The Register

© 2008 by The USENIX Association

All Rights Reserved

This volume is published as a collective work. Rights to individual papers remain with the author or the author's employer. Permission is granted for the noncommercial reproduction of the complete work for educational or research purposes. USENIX acknowledges all trademarks herein.

**8th USENIX Symposium on Operating Systems
Design and Implementation (OSDI '08)
December 8–10, 2008
San Diego, CA, USA**

Index of Authors	vi
Message from the Program Co-Chairs.	vii

Monday, December 8

Cloud Computing

DryadLINQ: A System for General-Purpose Distributed Data-Parallel Computing Using a High-Level Language . . 1 <i>Yuan Yu, Michael Isard, Dennis Fetterly, and Mihai Budiu, Microsoft Research Silicon Valley; Úlfar Erlingsson, Reykjavík University, Iceland, and Microsoft Research Silicon Valley; Pradeep Kumar Gunda and Jon Currey, Microsoft Research Silicon Valley</i>	
Everest: Scaling Down Peak Loads Through I/O Off-Loading. 15 <i>Dushyanth Narayanan, Austin Donnelly, Eno Thereska, Sameh Elnikety, and Antony Rowstron, Microsoft Research Cambridge, United Kingdom</i>	
Improving MapReduce Performance in Heterogeneous Environments 29 <i>Matei Zaharia, Andy Konwinski, Anthony D. Joseph, Randy Katz, and Ion Stoica, University of California, Berkeley</i>	

OS Architecture

Corey: An Operating System for Many Cores. 43 <i>Silas Boyd-Wickizer, Massachusetts Institute of Technology; Haibo Chen, Rong Chen, and Yandong Mao, Fudan University; Frans Kaashoek, Robert Morris, and Aleksey Pesterev, Massachusetts Institute of Technology; Lex Stein and Ming Wu, Microsoft Research Asia; Yuehua Dai, Xi'an Jiaotong University; Yang Zhang, Massachusetts Institute of Technology; Zheng Zhang, Microsoft Research Asia</i>	
CuriOS: Improving Reliability through Operating System Structure. 59 <i>Francis M. David, Ellick M. Chan, Jeffrey C. Carlyle, and Roy H. Campbell, University of Illinois at Urbana-Champaign</i>	
Redline: First Class Support for Interactivity in Commodity Operating Systems 73 <i>Ting Yang, Tongping Liu, and Emery D. Berger, University of Massachusetts Amherst; Scott F. Kaplan, Amherst College; J. Eliot B. Moss, University of Massachusetts Amherst</i>	

Monitoring

Network Imprecision: A New Consistency Metric for Scalable Monitoring. 87 <i>Navendu Jain, Microsoft Research; Prince Mahajan and Dmitry Kit, University of Texas at Austin; Praveen Yalagandula, HP Labs; Mike Dahlin and Yin Zhang, University of Texas at Austin</i>	
Lightweight, High-Resolution Monitoring for Troubleshooting Production Systems 103 <i>Sapan Bhatia, Princeton University; Abhishek Kumar, Google Inc.; Marc E. Fiuczynski and Larry Peterson, Princeton University</i>	
Automating Network Application Dependency Discovery: Experiences, Limitations, and New Solutions 117 <i>Xu Chen, University of Michigan; Ming Zhang, Microsoft Research; Z. Morley Mao, University of Michigan; Paramvir Bahl, Microsoft Research</i>	

Tuesday, December 9

File Systems

- SQCK: A Declarative File System Checker. 131
Haryadi S. Gunawi, Abhishek Rajimwale, Andrea C. Arpaci-Dusseau, and Remzi H. Arpaci-Dusseau, University of Wisconsin, Madison
- Transactional Flash. 147
Vijayan Prabhakaran, Thomas L. Rodeheffer, and Lidong Zhou, Microsoft Research, Silicon Valley
- Avoiding File System Micromanagement with Range Writes. 161
Ashok Anand and Sayandeep Sen, University of Wisconsin, Madison; Andrew Krioukov, University of California, Berkeley; Florentina Popovici, Google; Aditya Akella, Andrea Arpaci-Dusseau, Remzi Arpaci-Dusseau, and Suman Banerjee, University of Wisconsin, Madison

Programming Language Techniques

- Binary Translation Using Peephole Superoptimizers 177
Sorav Bansal and Alex Aiken, Stanford University
- R2: An Application-Level Kernel for Record and Replay. 193
Zhenyu Guo, Microsoft Research Asia; Xi Wang, Tsinghua University; Jian Tang and Xuezheng Liu, Microsoft Research Asia; Zhilei Xu, Tsinghua University; Ming Wu, Microsoft Research Asia; M. Frans Kaashoek, MIT CSAIL; Zheng Zhang, Microsoft Research Asia
- KLEE: Unassisted and Automatic Generation of High-Coverage Tests for Complex Systems Programs. 209
Cristian Cadar, Daniel Dunbar, and Dawson Engler, Stanford University

Security

- Hardware Enforcement of Application Security Policies Using Tagged Memory 225
Nickolai Zeldovich, Massachusetts Institute of Technology; Hari Kannan, Michael Dalton, and Christos Kozyrakis, Stanford University
- Device Driver Safety Through a Reference Validation Mechanism 241
Dan Williams, Patrick Reynolds, Kevin Walsh, Emin Gün Sirer, and Fred B. Schneider, Cornell University
- Digging for Data Structures. 255
Anthony Cozzie, Frank Stratton, Hui Xue, and Samuel T. King, University of Illinois at Urbana-Champaign

Dealing with Concurrency Bugs

- Finding and Reproducing Heisenbugs in Concurrent Programs 267
Madanlal Musuvathi, Shaz Qadeer, and Thomas Ball, Microsoft Research; Gerard Basler, ETH Zurich; Piramanayagam Arumuga Nainar, University of Wisconsin, Madison; Iulian Neamtiu, University of California, Riverside
- Gadara: Dynamic Deadlock Avoidance for Multithreaded Programs. 281
Yin Wang, University of Michigan and Hewlett-Packard Laboratories; Terence Kelly, Hewlett-Packard Laboratories; Manjunath Kudlur, Stéphane Lafortune, and Scott Mahlke, University of Michigan
- Deadlock Immunity: Enabling Systems to Defend Against Deadlocks 295
Horatiu Julia, Daniel Tralamazza, Cristian Zamfir, and George Candea, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

Wednesday, December 10

Various Good Things

Difference Engine: Harnessing Memory Redundancy in Virtual Machines 309
Diwaker Gupta, University of California, San Diego; Sangmin Lee, University of Texas at Austin; Michael Vrable, Stefan Savage, Alex C. Snoeren, George Varghese, Geoffrey M. Voelker, and Amin Vahdat, University of California, San Diego

Quanto: Tracking Energy in Networked Embedded Systems 323
Rodrigo Fonseca, University of California, Berkeley, and Yahoo! Research; Prabal Dutta, University of California, Berkeley; Philip Levis, Stanford University; Ion Stoica, University of California, Berkeley

Leveraging Legacy Code to Deploy Desktop Applications on the Web 339
John R. Douceur, Jeremy Elson, Jon Howell, and Jacob R. Lorch, Microsoft Research

Wide-Area Distributed Systems

FlightPath: Obedience vs. Choice in Cooperative Services. 355
Harry C. Li and Allen Clement, University of Texas at Austin; Mirco Marchetti, University of Modena and Reggio Emilia; Manos Kapritsos, Luke Robison, Lorenzo Alvisi, and Mike Dahlin, University of Texas at Austin

Mencius: Building Efficient Replicated State Machine for WANs 369
Yanhua Mao, CSE, University of California, San Diego; Flavio P. Junqueira, Yahoo! Research Barcelona; Keith Marzullo, CSE, University of California, San Diego