Dear reader,

Welcome to the Proceedings of the 22nd ACM Symposium on Operating Systems Principles–SOSP’09, held in Big Sky, Montana. The 23 papers published here explore a wide range of computer systems topics, from operating system device driver design to system designs for improved security and scalability. Collectively these papers report on some of the most creative and thought-provoking ideas in computer systems today. The papers illustrate the increasing international character of operating research, with co-authors from institutions in 8 different countries represented among the 23 accepted papers. All accepted papers were shepherded by PC members to ensure that they are as complete and easy to read as possible. We hope you will enjoy these papers as much as we did in selecting them.

Foremost, we would like to thank the authors of the 139 submitted papers, both those selected and those not, for choosing SOSP as the venue to submit their work. The success of the conference rests on the effort and creativity of its authors. Selecting 23 papers out of the set of submitted papers was difficult because so many of the submissions were of high quality. To make the selection process as fair and as consistent as possible, we used the same reviewing structure as SOSP’07 and other recent systems conferences, with 13 “heavy”-load and 20 “light”-load program committee members. The heavy-load members reviewed about 35 submissions each and attended the face-to-face PC meeting in Seattle, Washington. The light-load members reviewed about 15 papers each and did not attend the PC meeting. With a three-stage reviewing process involving 758 separate reviews, all papers following the submission guidelines received at least three reviews, and some papers received as many as ten. 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, 10 of the accepted papers were co-authored by PC members. We would like to take this opportunity to thank the program committee members, and the external reviewers, for their dedication and hard work.

We introduced one innovation in the program committee discussion. Based on a recent suggestion made by JD Douceur, we asked each PC member to rank order the papers they reviewed, prior to the PC meeting. We then computed a global order that (approximately) minimized the aggregate mean square error of the global order with respect to the individual rankings of the PC members. Ranking allowed the PC meeting to focus on those papers for which we lacked consensus, and most of the PC felt the ranking data was helpful. The calculated global rank in the end predicted 17 of the 23 papers that were selected for the conference. We notified selected authors as to the rank of their papers, but we chose not to publish the rank for wider dissemination.

We collected some additional statistics of interest. Authors of selected papers reported that their projects took a median of four person-years of effort, suggesting that the typical SOSP paper represents ambitious, difficult work. Eight of the 23 selected papers were co-authored by a researcher at Microsoft, but it is also important to note that Microsoft co-authors submitted 31 papers. Authors could “opt-in” to allow their papers to be reviewed by a shadow PC consisting of graduate students at various schools. The shadow PC was run independently, and the regular PC was not allowed to see the shadow PC reviews. Nevertheless, papers where authors allowed shadow PC access, presumably reflecting that the authors believed the paper was ready for wider review, were roughly twice as likely to be selected for publication by the regular PC (17 accepted out of 81 marked as reviewable). In the other direction, a significant fraction of submitted papers (30) ignored the formatting
guidelines (e.g. length restrictions), presumably because of haste in preparing the paper under a deadline. Papers that ignored the formatting guidelines were roughly half as likely to be selected as those that followed the guidelines.

The Women’s Workshop organized for the 20th anniversary of the SYSTERS group and held before SOSP 2007 was a great success. Building on this foundation, a new workshop, Diversity, was organized and co-located with OSDI 2008. Diversity expands on the outreach and mentoring goals of the original Women’s Workshop by including topics of interest to both women and to underrepresented minorities. The second offering of this Diversity workshop will be held at this year’s SOSP, cementing what we hope will be an annual workshop alternating between OSDI and SOSP.

Diversity is one of nine co-located workshops at this SOSP. This will be the first SOSP to offer such a full slate of workshops. Six of the workshops will be held on the Saturday and Sunday before SOSP and three will be held after SOSP on Wednesday afternoon and evening. To maintain the single track tradition of SOSP, no workshop activities overlap with regular conference events. We would like to thank the organizers and sponsors of all nine workshops as well as Marc Fiuczynski, who served as our first Workshop Chair. We welcome community feedback on the addition of a workshop program to future SOSPs.

SOSP is a great conference mostly because it attracts so many high-quality submissions, and we would like to thank again all the authors who submitted. We also 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: Geoff Kuenning (Treasurer), Sharon Weber (Sponsorship), Jay Lorch (Registration), Dilma Da Silva (Publicity), Yvonne Coady (Scholarships), Mike Swift (Posters), Mike Dahlin (WIPs), Chris Small (BoFs), Marc Fiuczynski (Workshops), Ken Birman (Hall of Fame Awards), Todd Deshane (Conference Webmaster) and John P. John (Submission Webmaster). 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.

Jeanna Matthews  
General Chair  
Clarkson University

Thomas Anderson  
Program Chair  
University of Washington
Table of Contents

22nd ACM Symposium on Operating Systems Principles Organization .......... vii

Keynote Address: ACM Turing Award Lecture
Barbara Liskov (Massachusetts Institute of Technology)

Session 1: Scalability
Session Chair: Jeff Dean (Google)

- FAWN: A Fast Array of Wimpy Nodes ................................................................. 1
  David G. Andersen, Jason Franklin (Carnegie Mellon University), Michael Kaminsky (Intel Labs),
  Amar Phanishayee, Lawrence Tan, Vijay Vasudevan (Carnegie Mellon University)

- RouteBricks: Exploiting Parallelism To Scale Software Routers .................. 15
  Mihai Dobrescu (EPFL), Norbert Egi (Lancaster University), Katerina Argyraki (EPFL),
  Byung-Gon Chun, Kevin Fall, Gianluca Iannaccone, Allan Knies, Maziar Manesh, Sylvia Ratnasamy
  (Intel Research Laboratories)

- The Multikernel: A New OS Architecture for Scalable Multicore Systems ...... 29
  Andrew Baumann (ETH Zurich), Paul Barham (Microsoft Research, Cambridge),
  Pierre-Evariste Dagand (ENS Cachan Bretagne), Tim Harris, Rebecca Isaacs (Microsoft Research, Cambridge),
  Simon Peter, Timothy Roscoe, Adrian Schüpbach, Akhilesh Singhania (ETH Zurich)

Session 2: Device Drivers
Session Chair: Philip Levis (Stanford University)

- Fast Byte-Granularity Software Fault Isolation .................................................. 45
  Miguel Castro, Manuel Costa, Jean-Philippe Martin, Marcus Peinado, Periklis Akritidis, Austin Donnelly,
  Paul Barham, Richard Black (Microsoft Research, Cambridge)

- Tolerating Hardware Device Failures in Software ............................................ 59
  Asim Kadav, Matthew J. Renzelmann, Michael M. Swift (University of Wisconsin-Madison)

- Automatic Device Driver Synthesis with Termite .......................................... 73
  Leonid Ryzhyk, Peter Chubb, Ihor Kuz, Etienne Le Sueur (NICTA, The University of New South Wales),
  Gernot Heiser (NICTA, The University of New South Wales, Open Kernel Laboratories)

Session 3: Debugging
Session Chair: Yuanyuan Zhou (University of California, San Diego)

- Automatically Patching Errors in Deployed Software ........................................ 87
  Jeff H. Perkins (Massachusetts Institute of Technology), Sunghun Kim (HKUST), Sam Larsen (VMWare),
  Saman Amarasinghe, Jonathan Bachrach, Michael Carbin (Massachusetts Institute of Technology),
  Carlos Pacheco (BCG), Frank Sherwood (None), Stelios Sidiropoulos (Massachusetts Institute of Technology),
  Greg Sullivan (BAE), Weng-Fai Wong (NUS), Yoav Zibin (Come2Play),
  Michael D. Ernst (University of Washington), Martin Rinard (Massachusetts Institute of Technology)

- Debugging in the (Very) Large: Ten Years of Implementation and Experience ...... 103
  Kirk Glerum, Kinshuman Kinshuman, Steve Greenberg, Gabriel Aul, Vince Orgovan, Greg Nichols,
  David Grant, Gretchen Lohle, Galen Hunt (Microsoft Corporation)

- Detecting Large-Scale System Problems by Mining Console Logs ..................... 117
  Wei Xu (University of California, Berkeley), Ling Huang (Intel Labs Berkeley), Armando Fox, David Patterson,
  Michael I. Jordan (University of California, Berkeley)

Session 4: I/O
Session Chair: Peter Chen (University of Michigan)

- Better I/O Through Byte-Addressable, Persistent Memory ................................. 133
  Jeremy Condit, Edmund B. Nightingale (Microsoft Research),
  Christopher Frost (University of California, Los Angeles),
  Engin Ipek, Benjamin Lee, Doug Burger, Derrick Coetzee (Microsoft Research)
• **Modular Data Storage with Anvil** .......................................................... 147
  Mike Mammarella, Shant Hovsepian (*University of California, Los Angeles*),
  Eddie Kohler (*University of California, Los Angeles & Meraki*)

• **Operating Systems Transactions** ....................................................... 161
  Donald E. Porter, Owen S. Hofmann, Christopher J. Rossbach, Alexander Benn, Emmett Witchel
  (*The University of Texas at Austin*)

**Session 5: Parallel Debugging**
Session Chair: Jeff Chase (*Duke University*)

• **PRES: Probabilistic Replay with Execution Sketching on Multiprocessors** .................. 177
  Soyeon Park, Yuanjuan Zhou (*University of California, San Diego*), Weiwei Xiong, Zuoning Yin, Rini Kaushik,
  Kyu H. Lee, Shan Lu (*University of Illinois at Urbana Champaign*)

• **ODR: Output-Deterministic Replay for Multicore Debugging** ................................ 193
  Gautam Altekar, Ion Stoica (*University of California, Berkeley*)

**Session 6: Kernels**
Session Chair: Timothy Roscoe (*ETH Zurich*)

• **seL4: Formal Verification of an OS Kernel** ........................................ 207
  Gerwin Klein, Kevin Elphinstone (*NICTA & UNSW*), Gernot Heiser (*NICTA, UNSW & Open Kernel Labs*),
  June Andronick (*NICTA & UNSW*), David Cock, Philip Derrin (*NICTA*),
  Dhammika Elkaduwe (*NICTA & UNSW*), Kai Engelhardt (*UNSW & NICTA*), Rafal Kolanski (*NICTA & UNSW*),
  Michael Norrish (*NICTA & ANU*), Thomas Sewell (*NICTA*), Harvey Tuch, Simon Winwood (*NICTA & UNSW*)

• **Helios: Heterogeneous Multiprocessing with Satellite Kernels** ............................. 221
  Edmund B. Nightingale, Orion Hodson (*Microsoft Research*), Ross McLroy (*University of Glasgow*),
  Chris Hawblitzel, Galen Hunt (*Microsoft Research*)

• **Surviving Sensor Network Software Faults** ........................................... 235
  Yang Chen (*University of Utah*), Omprakash Gnawali (*University of Southern California*),
  Maria Kazandjieva, Philip Levis (*Stanford University*), John Regehr (*University of Utah*)

**Session 7: Clusters**
Session Chair: David Anderson (*Carnegie Mellon University*)

• **Distributed Aggregation for Data-Parallel Computing: Interfaces and Implementations** ................................................. 247
  Yuan Yu, Pradeep Kumar Gunda, Michael Isard (*Microsoft Research*)

• **Quincy: Fair Scheduling for Distributed Computing Clusters** .............................. 261
  Michael Isard, Vijayan Prabhakaran, Jon Currey, Udi Wieder, Kunal Talwar, Andrew Goldberg
  (*Microsoft Research*)

• **UpRight Cluster Services** ........................................................................ 277
  Allen Clement, Manos Kapritsos, Sangmin Lee, Yang Wang, Lorenzo Alvisi, Mike Dahlin, Taylor Riché
  (*The University of Texas at Austin*)

**Session 8: Security**
Session Chair: Stefan Savage (*University of California, San Diego*)

• **Improving Application Security with Data Flow Assertions** .............................. 291
  Alexander Yip, Xi Wang, Nickolai Zeldovich, M. Frans Kaashoek (*Massachusetts Institute of Technology*)

• **Heat-ray: Combating *Identity Snowball* Attacks Using Machine Learning, Combinatorial Optimization and Attack Graphs** .................................................. 305
  John Dunagan, Alice X. Zheng (*Microsoft Research*), Daniel R. Simon (*Microsoft*)

• **Fabric: A Platform for Secure Distributed Computation and Storage** .................... 321
  Jed Liu, Michael D. George, K. Vikram, Xin Qi, Lucas Waye, Andrew C. Myers (*Cornell University*)

**Author Index** ........................................................................................................... 335
22^{nd} ACM Symposium on Operating Systems Principles
Organization

General Chair: Jeanna Neefe Matthews (Clarkson University)
Program Chair: Thomas Anderson (University of Washington)
Treasurer: Geoff Kuenning (Harvey Mudd College)
Financial Support Chair: Sharon Weber (VMware)
Registration Chair: Jacob R. Lorch (Microsoft Research)
Publicity Chair: Dilma Da Silva (IBM Research)
Workshops Chair: Marc Fiuczynksi (Princeton University)
Scholarship Chair: Yvonne Coady (University of Victoria)
Poster Session Chair: Michael Swift (University of Wisconsin, Madison)
Work-in-progress Session Chair: Michael Dahlin (University of Texas at Austin)
Birds-of-a-Feather Chair: Christopher Small (BBN Technologies)
Hall of Fame Award Committee Chair: Ken Birman (Cornell University)
Conference Webmaster: Todd Deshane (Clarkson University)
Submission Webmaster: John P. John (University of Washington)

Program Committee: Lorenzo Alvisi (University of Texas at Austin)
                     David Andersen (Carnegie-Mellon University)
                     Thomas Anderson (University of Washington)
                     Andrea Arpaci-Dusseau (University of Wisconsin)
                     Paul Barham (Microsoft Research Cambridge)
                     Herbert Bos (Vrije Universiteit Amsterdam)
                     Mike Burrows (Google)
                     Miguel Castro (Microsoft Research Cambridge)
                     Jeff Chase (Duke University)
                     Peter Chen (University of Michigan)
                     Jon Crowcroft (Cambridge University)
                     Michael Dahlin (University of Texas at Austin)
                     Jeff Dean (Google)
                     Mike Freedman (Princeton University)
                     Hermann Härtig (TU Dresden)
                     Steven Hand (Cambridge University)
Program Committee (continued):  
Norm Hutchinson (University of British Columbia)  
Arvind Krishnamurthy (University of Washington)  
Max Krohn (Yale University)  
Philip Levis (Stanford University)  
Jeff Mogul (HP Labs)  
Andrew Myers (Cornell University)  
John Ousterhout (Stanford University)  
Sylvia Ratnasamy (Intel Berkeley)  
Timothy Roscoe (ETH Zurich)  
Stefan Savage (UC San Diego)  
Srini Seshan (Carnegie-Mellon University)  
Gun Sirer (Cornell University)  
Michael Swift (University of Washington)  
Helen Wang (Microsoft Research Redmond)  
Nickolai Zeldovich (MIT)  
Yuanyuan Zhou (University of Illinois)  
Zheng Zhang (Microsoft Research Asia)  

Additional reviewers:  
Bill Bolosky  
Allen Clement  
Jason Flinn  
Dan Grossman  
Tim Harris  
James Hendricks  
Harry Li  
Xuezheng Liu  
Jacob Lorch  
Dave Maltz  
Mirco Marchetti  
Frank McSherry  
Derek Murray  
Matthew Renzelmann  
Taylor Riché  
Henry Robinson  
Margo Seltzer  
Siddhartha Sen  
Haris Volos  
Emmett Witchell