



OPERATING SYSTEMS REVIEW

A Publication of the
Association for Computing Machinery
Special Interest Group on Operating Systems

Volume 35, Number 5

December, 2001

Proceedings of the

18th ACM Symposium on Operating Systems Principles (SOSP'01)

October 21-24, 2001

Chateau Lake Louise, Banff, Alberta, Canada

OPERATING SYSTEMS REVIEW

A Publication of the ACM Special Interest Group on Operating Systems

Chairperson

William E. Weihl
Akamai Technologies, Inc.
1400 Fashion Island Blvd.
San Mateo, CA 94404 USA
+1-650-627-5259
(bweihl@akamai.com)

Newsletter Editor

William M. Waite
Dept. of Electrical and
Computer Engineering
University of Colorado
Boulder, CO 80309-0425 USA
+1-303-492-7204
(William.Waite@Colorado.edu)

V. Chairperson

Valerie Issarny
INRIA - Rennes
IRISA
Campus de Beaulieu
35042 Rennes Cedex France
+33-2-99-84-7462
(issarny@irisa.fr)

Secretary-Treasurer

David Kotz
Dept. of Computer Science
Dartmouth College
6211 Sudikoff Laboratory
Hanover, NH 03755 USA
+1-603-646-1439
(dfk@cs.dartmouth.edu)

Information Director

Michael Dahlin
Department of Computer Sciences
Campus Mail Code: C0500
University of Texas
Austin, TX 78712 USA
+1-512-471-9549
(dahlin@cs.utexas.edu)

OPERATING SYSTEMS REVIEW is an informal publication of the ACM Special Interest Group on Operating Systems (SIGOPS), whose scope of interest includes: Computer operating systems and architecture for multiprogramming, multiprocessing, and time sharing; resource management; evaluation and simulation; reliability, integrity, and security of data; communications among computing processors; and computer system modeling analysis.

Membership in SIGOPS (at \$15 per annum) is open to ACM members and associate members, and student members (at \$8 per annum); non-ACM members may join also (at \$42 per annum). Non-members of SIGOPS may subscribe to OPERATING SYSTEMS REVIEW at \$30 per year. All SIGOPS members receive OPERATING SYSTEMS REVIEW.

SIGOPS membership application forms are available from ACM Headquarters (or see back cover of this issue), 1515 Broadway, 17th Floor, New York, New York 10036, telephone (212) 869-7440. Changes of address and other matters pertaining to the SIGOPS mailing list should be directed to ACM Headquarters, not to any of the SIGOPS officers.

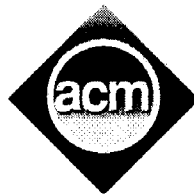
Contributions to OSR are sent to the editor and should be single spaced and printed on one side of the paper only. Text and diagrams must be contained within a 6.5 inch wide by 8.5 inch high (165mm by 216mm) area with at least one inch (25mm) top and side margins, 1.5 inch (38mm) bottom margin, and be camera-ready (i.e., be an original having sufficient contrast and density for reproduction). All letters to the editor will be considered for publication unless accompanied by a request to the contrary. Except for editorial items, all sources of material appearing in OPERATING SYSTEMS REVIEW will be clearly identified. Items attributed to individuals will ordinarily be interpreted as personal rather than organizational opinions. The technical material in this newsletter is not refereed. **Deadlines for submitting material are the 15th of February, May, August, and November. The SIGOPS conference proceedings will be mailed as the December issue in odd numbered calendar years, and the ASPLOS conference proceedings will be mailed in even numbered calendar years.**

OPERATING SYSTEMS REVIEW (ISSN 0163-5980) is published five times a year (January, April, July, October and December) by the Association for Computing Machinery Inc., 1515 Broadway, New York, NY 10036. Periodicals postage paid at New York, NY 10001, and at additional mailing offices. POSTMASTER: Send address changes to OPERATING SYSTEMS REVIEW, ACM, 1515 Broadway, New York, NY 10036.

Subscriptions: Annual subscription cost of \$12.53 is included in the member dues of \$15.00 (for students cost is included in \$8.00); the non-member annual subscription is \$42.00.

Proceedings of the
18th ACM Symposium on
Operating Systems Principles
(SOSP'01)

*October 21-24, 2001
Chateau Lake Louise, Banff, Alberta, Canada*



Sponsored by ACM SIGOPS
(Association for Computing Machinery Special Interest Group in Operating Systems).

Supported by Microsoft Research, Intel Corporation, QUALCOMM, Mercury Computer Systems,
Hewlett-Packard Labs, Akamai, and Inktomi.

Microsoft
Research

intel.[®]

QUALCOMM[®]

Computer Systems, Inc.
MERCURY



**The Association for Computing Machinery
1515 Broadway
New York, N.Y. 10036**

Copyright © 2001 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 the 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 E-mail 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, +1-978-750-8400, +1-978-750-4470 (fax).

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 was 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.

ACM ISBN: 1-58113-389-8

Additional copies may be ordered prepaid from:

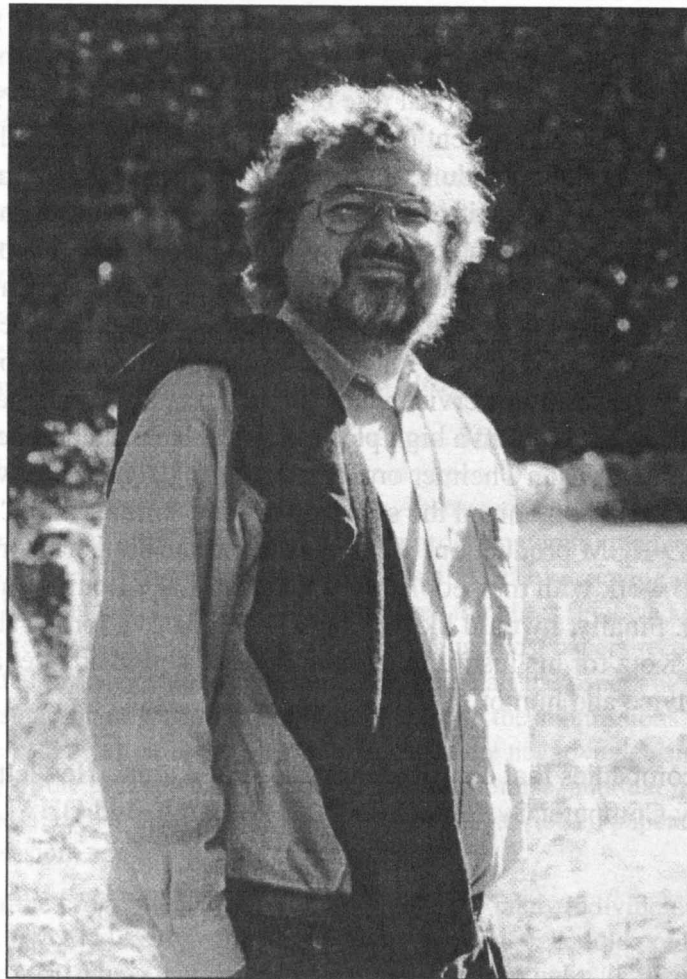
ACM Order Department
P.O. Box 11405
New York, N.Y. 10286-1405

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

ACM Order Number: 534010

Printed in the U.S.A.

**DEDICATED TO THE MEMORY OF
JOCHEN LIEDTKE
(1953-2001)**



Jochen Liedtke, a program committee member of this SOSP, died tragically at Frankfurt Airport on Sunday, June 10, 2001 on his return from the SOSP program committee meeting. He had been in ill health earlier in the year, and had undergone two major surgeries. It would have been easy (and entirely understandable) for him to have declined to travel to the meeting. Yet, he was committed to participating fully in the process and travelled across the Atlantic to attend. His dedication and sense of duty are truly humbling.

At the time of his death, Jochen was a professor at the University of Karlsruhe in Germany. Earlier, he had been a researcher at the IBM T.J. Watson Research Center at Yorktown Heights and at the GMD in Bonn. He is widely known in the operating systems community for his contributions to microkernel design and implementation, particularly the L3 and L4 microkernels. He was an active participant in the research community, serving on the program committees of many conferences and workshops. He served as the General Chair of HotOS-VIII, held in Schloss Elmau, Germany in May, 2001.

Jochen is survived by his wife, Adelheid, and many friends and colleagues. The SOSP community is deeply saddened by his passing, and dedicates these proceedings to his memory.

Eighteenth ACM Symposium on Operating System Principles

MESSAGE FROM THE GENERAL CHAIR

SOSP has moved back into the mountains, this time to the Chateau Lake Louise in Alberta.

The job of a General Chair is, basically, to worry. The people behind this SOSP were outstanding and gave me little occasion to exercise the traditional role of a General Chair. Satya's vision and hard work with his committee has resulted in an excellent program that is both broad and deep. Tom Bressoud as treasurer made sure that the financial procedures – from contract negotiation to registration cash flow – worked smoothly, and was always ready with the right spreadsheet when I felt my worry level rising. Norm Hutchinson is the main reason that the conference is in this beautiful historic hotel and in this stunning location, and he has managed the logistics wonderfully. Lorenzo Alvisi ran publicity smoothly despite the deep nondeterministic mysteries of posting to netnews and the interesting cyber-counterattacks of those who felt e-mail containing a call for papers were the vilest form of spam. Geoff Voelker built and maintained the web registration service while simultaneously serving as SIGCOMM local arrangements chair, which should earn him a high place in ACM Heaven, and he kept us all sane in tense moments with his good humor. Marvin Theimer organized a strong committee which ably distributed the student scholarships, Mike Feeley organized the student help and scribes, and Chris Small once again arranged the conference CD-ROM production. We also enjoyed the support from many at ACM Headquarters, and it was a joy to work with the people at the Chateau Lake Louise, most notably Amanda Burrows and Joanne Trudeau. Finally, I would like to thank Ava Hristova for her design of the conference web page structure, David Kotz for his institutional memory and sound advice, and Stefan Savage and Fred Schneider for their advice and humor.

I am very thankful to the companies that contributed to SOSP: Akamai, Hewlett-Packard Labs, Inktomi, Intel Corporation, Mercury Computer Systems, Microsoft Research, and QUALCOMM.

Thank you all!

Keith Marzullo
University of California, San Diego

Eighteenth ACM Symposium on Operating System Principles

MESSAGE FROM THE PROGRAM CHAIR

At the dawn of the new millennium, it is useful to pause and reflect on the role of SOSP. It is hard to believe that SOSP is over 35 years old! The first one was held in 1965, long before many of the attendees of this conference were born. In those 35 years, computing has been transformed in ways that were unimaginable back in 1965. Somehow, SOSP has retained its reputation as the premier forum for research results in experimental computer systems. How has it succeeded in this role for so long?

I believe that there are two reasons. First and foremost is the dedication and hard work of the many individuals who have served on the SOSP program committees. These are talented individuals with heavy demands on their time. Their willingness to devote a substantial part of that time to the SOSP review process shows their commitment to preserving the quality of SOSP. The second reason is SOSP's broad charter. In spite of the emergence of more specialized conferences in areas such as networking, mobile computing, reliability and security, SOSP continues to welcome work of the highest quality in those areas. This broad view of experimental computer systems has served SOSP well, and will continue to be one of its strengths as we discover further new dimensions to computing.

This year we received 85 submissions: lower than in the past two SOSP's, but well within the observed long-term range. We accepted 17 of the 85, and nominated two as award papers to ACM TOCS ("Untrusted Hosts and Confidentiality: Secure Program Partitioning" and "BASE: Using Abstraction to Improve Fault Tolerance"). All acceptances were conditional, pending successful shepherding by a program committee member. This long-standing SOSP tradition has been an important factor in ensuring high quality. The reviewing process we used was a variant of one that has worked well in the past. Four program committee members reviewed each paper; each of those members also obtained one review from an external reviewer. Thus, most papers received eight reviews. In addition, each program committee member was required to skim about half the papers he or she did not review. This combination of deep understanding as well as broad familiarity with the submissions led to thoughtful and well-informed discussions at the program committee meeting. Some of the negative decisions were painful and heart-breaking: we felt that many of the rejected papers showed a lot of promise, and that improved versions would be strong future submissions to SOSP or other conferences. Besides the reviewed papers, the conference has three invited speakers, a poster session, and a work-in-progress session.

One of the pleasant tasks of the program chair is to acknowledge the many individuals who have helped in putting the program together. First, I would like to thank the program committee. I could not have asked for a better group of people to work with. They did a huge amount of reviewing, were prompt with their reviews, participated vigorously in the discussions at the meeting, were fair and balanced in their judgments, and kept an open mind at all times. Whenever I needed advice on how to handle a tricky situation, they came through with thoughtful suggestions and carefully considered opinions. Overall, they were superb custodians of SOSP quality.

I would also like to thank the Publications Chair, Greg Ganger, and his assistant, Joan Digney, for taking care of all aspects of producing this document. Their help has saved me a considerable amount of time and attention to detail. Chris Small handled the production of a CD-ROM companion to the proceedings, and I would also like to thank him. Keith Marzullo has been wonderful to work with as General Chair: prompt in his responses, thorough and detailed in his follow up, and always in good spirits. To my administrative assistant, Tracy Farbacher, and my graduate student, Rajesh Balan, both of whom helped me in countless ways with the logistics of the reviewing process and the program committee meeting: a big thank you! The dedication and help of all these people, and many others not mentioned here, has made it a pleasure to serve as program chair of this SOSP. It has also been an honor and privilege.

M. Satyanarayanan
Program Chair, SOSP-18
August 2001

Eighteenth ACM Symposium on Operating System Principles

EXTERNAL REVIEWERS

Ali-Reza Adl-Tabatabai	John Douceur	Peter Honeyman
Atul Adya	Fred Douglass	Ken Hopkinson
Bikash Agarwalla	Richard Draves	John Howard
Adnan Agbaria	Peter Druschel	Liviu Iftode
Marcos Aguilera	Dan Duchamp	Rebecca Isaacs
Ahmed Amer	Kenneth Duda	Sitaram Iyer
David Andersen	Derek Eager	Kyle Jamieson
Andrea Arpaci-Dusseau	Keith Edwards	John Jannotti
Uwe Assmann	Greg Eisenhauer	Minwen Ji
Benjamin Atkin	Daniel Ellard	David Johnson
Mary Baker	Carla Ellis	Alan Jones
Magdalena Balazinska	Carl Ellison	Anthony Joseph
Gaurav Banga	Kevin Elphinstone	Frans Kaashoek
Paul Barham	Dawson Engler	Scott Karlin
Rimon Barr	Michael Ernst	Terence Kelly
Tugkan Batu	Deborah Estrin	Minkyong Kim
Andy Bavier	Kevin Fall	Jay Kistler
Frank Bellosa	Anton Faradjian	Emil Kohn
Chuck Blake	Keith Farkas	Olga Kornievskaia
Bill Bolosky	Nick Feamster	Lakshman Krishnamurthy
Dan Boneh	Jason Flinn	Arvind Krishnamurthy
Gaetano Borriello	Liana Fong	Geoffrey Kuenning
Aaron Brown	Josh Fryman	Sanjeev Kumar
Barry Brumitt	Ayalvadi Ganesh	Yui-Wah Lee
Emma Brunskill	Ivan Ganev	Kai Li
Ken Calvert	Greg Ganger	Jorg Liebeherr
Andrew Campbell	Ada Gavrilovska	Gerd Lieflaender
Miguel Castro	Sanjay Ghemawat	Ada Lim
Ugur Cetintemel	Vadim Gleizer	Jack Lo
Fay Chang	Yitzchak Gottlieb	Darrell Long
Benjie Chen	Ramesh Govindan	David Lowenthal
Brad Chen	Paul Graunke	Honghui Lu
Peter Chen	Steven Gribble	Kostas Magoutis
Yuqun Chen	Robert Grimm	Scott Mainwaring
Tzi-cker Chiueh	Mark Gritter	Durga Devi Mannaru
Michal Cierniak	Dirk Grunwald	David Mazieres
Jamieson Cobleigh	Emin GunSirer	Patrick McDaniel
Andy Collins	Frank Hady	Frank Mehnert
Mark Corner	Wei Han	Mike Mesnier
Landon Cox	Steven Hand	Yaron Minsky
David Culler	Mor Harchol-Balter	Vishal Misra
Dorothy Curtis	Gernot Heiser	Allen Miu
Steven Czerwinski	Johannes Helander	Jeffrey Mogul
Uwe Dannowski	Michael Hohmuth	Robert Morris
Eyal de Lara	David Holland	Hans Mulder

Eighteenth ACM Symposium on Operating System Principles

EXTERNAL REVIEWERS, continued

Sape Mullender	Mendel Rosenblum	Michael Swift
David Nagle	Daniela Rosu	Renu Tewari
Erich Nahum	Marcel Rosu	Marvin Theimer
Aki Nakao	Dan Rubenstein	Walter Tichy
Dushyanth Narayanan	Algis Rudys	Michael Ubell
Badri Nath	Yasushi Saito	Volkmar Uhlig
Juan Navarro	Constantine Sapuntzakis	Theo Ungerer
Rich Neves	Stefan Saroiu	Amin Vaahdat
Tin-Fook Ngai	Rick Schlichting	Amin Vaahdat
Nathaniel Nystrom	Brian Schmidt	Robbert vanRenesse
James O'Toole	Sebastian Schoenberg	Raj Vaswani
Vivek Pai	Alper Sen	Vivekanand Vellanki
Henrique Paques	Mauricio Serrano	Roman Vitenberg
Yoonho Park	Srinivasan Seshan	Geoff Voelker
Jignesh Patel	Konstantin Shagin	Jim Waldo
Arnab Paul	Marc Shapiro	Carl Waldspurger
Vern Paxson	Liuba Shrira	Limin Wang
Trevor Pering	Dan Simon	Randy Wang
Karin Petersen	Ambuj Singh	Yi-Min Wang
Christian Poellabauer	Jaswinder Pal Singh	Roy Want
Atul Prakash	Espen Skoglund	Weldon Washburn
Niels Provos	Christopher Small	Matt Welsh
Calton Pu	Frederick Smith	Patrick Widener
Xiaohu Qie	Keith Smith	John Wilkes
Lili Qiu	Alex Snoeren	Alec Wolman
Raj Rajkumar	Tammo Spalink	Theodore Wong
Suchitra Raman	Oliver Spatscheck	Rumi Zahir
Valentin Razmov	Christopher Stein	Steve Zdancewic
John Regehr	Daniel Stodolsky	Matthew Zekauskas
John Reumann	Ion Stoica	Ming Zhang
Lars Reuther	Daniel Sturman	Lantian Zheng
Martin Rinard	Sree Subramoney	Dong Zhou
Rodrigo Rodrigues	David Sullivan	Lidong Zhou
Timothy Roscoe	Florin Sultan	

Eighteenth ACM Symposium on Operating System Principles

CONFERENCE OFFICERS

General Chair:	Keith Marzullo, <i>UC San Diego</i>
Program Chair:	M. Satyanarayanan, <i>Carnegie Mellon University and Intel Research Pittsburgh</i>
Program Committee:	Hari Balakrishnan, <i>MIT</i> Maurice Herlihy, <i>Brown University</i> Mike Jones, <i>Microsoft Research</i> Monica Lam, <i>Stanford University</i> Butler Lampson, <i>Microsoft Research</i> Jochen Liedtke, <i>University of Karlsruhe</i> Andrew Myers, <i>Cornell University</i> Brian Noble, <i>University of Michigan</i> Larry Peterson, <i>Princeton University</i> Margo Seltzer, <i>Harvard University</i> David Tennenhouse, <i>Intel Research</i> John Zahorjan, <i>University of Washington</i>
Scholarship Chair:	Marvin Theimer, <i>Microsoft Research</i>
Scholarship Committee:	Cary Gray, <i>Wheaton College</i> Sharon Perl, <i>Compaq Systems Research Center</i> David Steere, <i>Oregon Graduate Institute</i> Werner Vogels, <i>Cornell University</i>
Local Arrangements:	Norm Hutchinson, <i>University of British Columbia</i>
Publication:	Greg Ganger, <i>Carnegie Mellon University</i>
Finance:	Tom Bressoud, <i>Lucent</i>
Registration:	Geoff Voelker, <i>UC San Diego</i>
Publicity:	Lorenzo Alvisi, <i>UT Austin</i>
Scribes and Volunteers:	Michael Feeley, <i>University of British Columbia</i>
CD ROM:	Christopher Small, <i>Sun Microsystems</i>

Eighteenth ACM Symposium on Operating System Principles

TABLE OF CONTENTS

Trust and Dependability

<i>Untrusted Hosts and Confidentiality: Secure Program Partitioning</i> Steve Zdancewic, Lantian Zheng, Nathaniel Nystrom, Andrew C. Myers (Cornell University).....	1
<i>BASE: Using Abstraction to Improve Fault Tolerance</i> Rodrigo Rodrigues, Barbara Liskov (MIT), Miguel Castro (Microsoft Research).....	15
<i>The Costs and Limits of Availability for Replicated Services</i> Haifeng Yu, Amin Vahdat (Duke University).....	29

Deconstructing the OS

<i>Information and Control in Gray-Box Systems</i> Andrea Arpaci-Dusseau, Remzi Arpaci-Dusseau (University of Wisconsin, Madison).....	43
<i>Bugs as Deviant Behavior: A General Approach to Inferring Errors in Systems Code</i> Dawson Engler, David Yu Chen, Seth Hallem, Andy Chou, Benjamin Chelf (Stanford University).....	57
<i>An Empirical Study of Operating Systems Errors</i> Andy Chou, Jun-Feng Yang, Benjamin Chelf, Seth Hallem, Dawson Engler (Stanford University)	73

Resource Management

<i>Real-Time Dynamic Voltage Scaling for Low-Power Embedded Operating Systems</i> Padmanabhan Pillai, Kang G. Shin (University of Michigan).....	89
<i>Managing Energy and Server Resources in Hosting Centres</i> Jeffrey S. Chase, Darrell C. Anderson, Prachi N. Thakar, Amin M. Vahdat (Duke University), Ronald P. Doyle (IBM Research Triangle Park).....	103
<i>Anticipatory Scheduling: A Disk Scheduling Framework to Overcome Deceptive Idleness in Synchronous I/O</i> Sitaram Iyer, Peter Druschel (Rice University).....	117

Networking

<i>Resilient Overlay Networks</i> David G. Andersen, Hari Balakrishnan, M. Frans Kaashoek, Robert Morris (MIT).....	131
<i>Building Efficient Wireless Sensor Networks with Low-Level Naming</i> John Heidemann, Fabio Silva, Chalermek Intanagonwivat, Ramesh Govindan (USC/ISI), Deborah Estrin, Deepak Ganesan (UCLA)	146
<i>Mesh Based Content Routing using XML</i> Alex C. Snoeren, Kenneth Conley, David K. Gifford (MIT)	160

Eighteenth ACM Symposium on Operating System Principles

TABLE OF CONTENTS, continued

File Systems

<i>A Low-Bandwidth Network File System</i> Athicha Muthitacharoen, Benjie Chen (MIT), David Mazières (NYU).....	174
<i>Storage Management and Caching in PAST, A Large-scale, Persistent Peer-to-peer Storage Utility</i> Antony Rowstron (Microsoft Research), Peter Druschel (Rice University)	188
<i>Wide-Area Cooperative Storage with CFS</i> Frank Dabek, M. Frans Kaashoek, David Karger, Robert Morris (MIT), Ion Stoica (UC Berkeley)	202

Event-Driven Architectures

<i>Building a Robust Software-Based Router Using Network Processors</i> Tammo Spalink, Scott Karlin, Larry Peterson, Yitzchak Gottlieb (Princeton University).....	216
<i>SEDA: An Architecture for Well-Conditioned, Scalable Internet Services</i> Matt Welsh, David Culler, Eric Brewer (UC Berkeley)	230

Eighteenth ACM Symposium on Operating System Principles

AUTHOR INDEX

Andersen, David G	131	Iyer, Sitaram.....	117
Anderson, Darrell C.	103	Kaashoek, M. Frans	131, 202
Arpaci-Dusseau, Andrea	43	Karger, David.....	202
Arpaci-Dusseau, Remzi.....	43	Karlin, Scott.....	216
Balakrishnan, Hari.....	131	Liskov, Barbara.....	15
Brewer, Eric	230	Mazières, David.....	174
Castro, Miguel.....	15	Morris, Robert	131, 202
Chase, Jeffrey S.....	103	Muthitacharoen, Athicha	174
Chelf, Benjamin	57, 73	Myers, Andrew C.	1
Chen, Benjie	174	Nystrom, Nathaniel.....	1
Chen, David Yu	57	Peterson, Larry.....	216
Chou, Andy	57, 73	Pillai, Padmanabhan	89
Conley, Kenneth.....	160	Rodrigues, Rodrigo.....	15
Culler, David	230	Rowstron, Antony.....	188
Dabek, Frank	202	Shin, Kang G.....	89
Doyle, Ronald P.....	103	Silva, Fabio	146
Druschel, Peter	117, 188	Snoeren, Alex C.....	160
Engler, Dawson	57, 73	Spalink, Tammo	216
Estrin, Deborah	146	Stoica, Ion	202
Ganesan, Deepak.....	146	Thakar, Prachi N.	103
Gifford, David K.	160	Vahdat, Amin M.	29, 103
Gottlieb, Yitzchak	216	Welsh, Matt.....	230
Govindan, Ramesh	146	Yang, Jun-Feng	73
Hallem, Seth.....	57, 73	Yu, Haifeng.....	29
Heidemann, John.....	146	Zdancewic, Steve.....	1
Intanagonwiwat, Chalermek	146	Zheng, Lantian.....	1