

October 5-7, 2015
Monterey, California, USA



Association for
Computing Machinery

Advancing Computing as a Science & Profession



SOSP'15

Proceedings of the Twenty-Fifth ACM

Symposium on Operating Systems Principles

Sponsored by:

ACM SIGOPS and UCSC SSRC

Supported by:

**Facebook, Google, HP, Intel, Microsoft Research,
Cisco, Datrium, VMware,**

Akamai, Bromium, Coho Data, Cloudera, IBM Research, Oracle



**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 © 2015 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: 978-1-4503-3834-9

Foreword

Dear Colleagues,

Welcome to the Proceedings of the 25th ACM Symposium on Operating Systems Principles (SOSP 2015), held at the Portola Hotel & Spa, Monterey, California, USA. This year's program includes 30 wonderful papers which cover topics ranging from high performance transaction processing systems to formal specification for practical implementations, from techniques for energy-aware systems to experiences with globally distributed software in the wild, from bug finding to big data. The program committee worked hard to select some of the most creative and interesting ideas in computer systems research today, and each accepted paper was lovingly shepherded by a program committee member to make sure the papers are as readable and complete as possible. We hope you will enjoy the program as much as we did in selecting it.

This "silver anniversary" edition of SOSP is being held along with an exciting collection of workshops. Of particular interest is a special History Day event which will feature talks from luminaries in the field including Jack Dennis, Peter Denning, Butler Lampson, Barbara Liskov, M. Satyanarayanan (Satya), Ken Birman, Andrew Herbert, Andrew Birrell, Dave Clark, Frans Kaashoek and Jeff Dean. We hope to make video recordings available online for posterity. Furthermore, we are pleased to report that this year's conference will continue the Open Access policy that debuted in 2013: all papers will be hosted on the ACM digital library, and available for viewing or download via a SIGOPS-provided link. Our thanks to the SIGOPS organizers and the community for setting this new standard.

This year we received a record-breaking 181 paper submissions which went forward for review in multiple rounds. The expert program committee was divided into 16 "heavy" and 14 "light" PC members, and in the first round we arranged for every paper to receive three reviews from PC members; approximately 50% of papers went through for a second round of PC reviewing, with a few additional expert reviews being contributed. Following this we had a short 3rd round of reviewing involving heavy PC members only. All in all, we produced 824 reviews, just 11 of which were from external experts. In addition to reviewing, we had a vibrant online discussion involving the entire PC which led to us selecting 78 papers to be discussed in person. The heavy PC members convened at a meeting hosted at Google in Mountain View, CA, where we selected 30 papers. All discussed papers received a description of the PC discussion. Throughout the process, author anonymity was maintained and conflicts of interest precluded; PC chair conflicts were dealt with in a parallel process run by Eddie Kohler. We hope that all authors found the reviews and discussion summaries helpful.

As always we are extremely thankful for our generous corporate sponsors. Their donations make it possible to host SOSPP in an environment that is conducive to collegial interaction. This year they have provided funds which have enabled us to award full or partial travel grants to over 100 students from a wide range of countries and institutions.

SOSP is a great conference mostly because it attracts so many high-quality submissions, and we would like to again thank *every* author for choosing to submit to SOSPP this year. We also thank the PC members for the tremendous amount of work they did: reviewing the submissions, providing feedback, and shepherding the accepted submissions. We are grateful to the external reviewers who provided an additional perspective on a few papers. SOSPP has always been organized by volunteer efforts from a host of people; we would like to thank all the following people who have dedicated so much of their time to the conference: Jacob Lorch (Treasurer), Kevin Greenan (local arrangements), Edouard Bugnion (sponsorships), Nitin Agrawal (scholarships), Joel and Shayna Frank (publicity), Peter Desnoyers (workshops), Daniel Peek (posters) and Andrew Birrell (publications).

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

Ethan Miller
General Chair

Steven Hand
Program Chair

Contents

Formal Systems

Chair: Petros Maniatis

IronFleet: Proving Practical Distributed Systems Correct	1
Chris Hawblitzel, Jon Howell, Manos Kapritsos, Jacob R. Lorch, Bryan Parno, Michael L. Roberts, Srinath Setty, Brian Zill (Microsoft Research)	
Using Crash Hoare Logic for Certifying the FSCQ File System	18
Haogang Chen, Daniel Ziegler, Tej Chajed, Adam Chlipala, M. Frans Kaashoek, Nickolai Zeldovich (MIT CSAIL)	
SibyIFS: formal specification and oracle-based testing for POSIX and real-world file systems	38
Tom Ridge (University of Leicester), David Sheets (University of Cambridge), Thomas Tuerk (FireEye), Andrea Giugliano (University of Leicester), Anil Madhavapeddy, Peter Sewell (University of Cambridge)	

Distributed Transactions

Chair: Tim Harris

No compromises: distributed transactions with consistency, availability, and performance	54
Aleksandar Dragojević, Dushyanth Narayanan, Edmund B Nightingale, Matthew Renzelmann, Alex Shamis, Anirudh Badam, Miguel Castro (Microsoft Research)	
Implementing Linearizability at Large Scale and Low Latency	71
Collin Lee, Seo Jin Park, Ankita Kejriwal (Stanford University), Satoshi Matsushita (NEC), John Ousterhout (Stanford University)	
Fast In-memory Transaction Processing using RDMA and HTM	87
Xingda Wei, Jiaxin Shi, Yanzhe Chen, Rong Chen, Haibo Chen (Shanghai Jiao Tong University)	

Distributed Systems

Chair: Marcos Aguilera

Paxos Made Transparent	105
Heming Cui (University of Hong Kong and Columbia University), Rui Gu, Cheng Liu (Columbia University), Tianyu Chen (Tsinghua University), Junfeng Yang (Columbia University)	
E2: A Framework for NFV Applications	121
Shoumik Palkar, Chang Lan, Sangjin Han (UC Berkeley), Keon Jang (Intel Labs), Aurojit Panda, Sylvia Ratnasamy (UC Berkeley), Luigi Rizzo (Università di Pisa), Scott Shenker (UC Berkeley and ICSI)	
Vuvuzela: Scalable Private Messaging Resistant to Traffic Analysis	137
Jelle van den Hooff, David Lazar, Matei Zaharia, Nickolai Zeldovich (MIT CSAIL)	

Concurrency and Performance

Chair: Emmett Witchel

Parallelizing User-Defined Aggregations using Symbolic Execution	153
Veselin Raychev (ETH Zurich), Madanlal Musuvathi, Todd Mytkowicz (Microsoft Research)	
Read-Log-Update: A Lightweight Synchronization Mechanism for Concurrent Programming	168
Alexander Matveev (MIT), Nir Shavit (MIT & Tel-Aviv University), Pascal Felber, Patrick Marlier (University of Neuchâtel)	
COZ: Finding Code that Counts with Causal Profiling	184
Charlie Curtsinger (Grinnell College), Emery D. Berger (University of Massachusetts Amherst)	

Energy Aware Systems

Chair: Landon Cox

JouleGuard: Energy Guarantees for Approximate Applications	198
Henry Hoffmann (University of Chicago)	
Software Defined Batteries	215
Anirudh Badam, Ranveer Chandra, Jon A. Dutra (Microsoft), Anthony Ferrese (Tesla Motors), Steve Hodges (Microsoft), Pan Hu (University of Massachusetts Amherst), Julia Meinershagen, Thomas Moscibroda, Bodhi Priyantha (Microsoft), Evangelia Skiani (Columbia Univeristy)	
Drowsy Power Management	230
Matthew Lentz, James Litton, Bobby Bhattacharjee (University of Maryland)	

More Distributed Transactions

Chair: Haibo Chen

- Yesquel: Scalable SQL storage for Web applications** 245
Marcos K. Aguilera (VMware Research Group), Joshua B. Leners (UT Austin, NYU), Michael Walfish (NYU)
- Building Consistent Transactions with Inconsistent Replication** 263
Irene Zhang, Naveen Kr. Sharma, Adriana Szekeres, Arvind Krishnamurthy, Dan R. K. Ports (University of Washington)
- High-Performance ACID via Modular Concurrency Control** 279
Chao Xie, Chunzhi Su, Cody Littlely, Lorenzo Alvisi (University of Texas at Austin), Manos Kapritsos (Microsoft Research), Yang Wang (Ohio State University)

Experience and Practice

Chair: Shan Lu

- Existential Consistency: Measuring and Understanding Consistency at Facebook** . 295
Haonan Lu (University of Southern California and Facebook), Kaushik Veeraraghavan, Philippe Ajoux, Jim Hunt, Yee Jiun Song, Wendy Tobagus, Sanjeev Kumar (Facebook), Wyatt Lloyd (University of Southern California and Facebook)
- Virtual CPU Validation** 311
Nadav Amit, Dan Tsafir, Assaf Schuster (Technion — Israel Institute of Technology), Ahmad Ayoub, Eran Shlomo (Intel Corporation)
- Holistic Configuration Management at Facebook**..... 328
Chunqiang Tang, Thawan Kooburat, Pradeep Venkatachalam, Akshay Chander, Zhe Wen, Aravind Narayanan, Patrick Dowell, Robert Karl (Facebook Inc.)

Bugs and Analysis

Chair: Brad Karp

- Failure Sketching: A Technique for Automated Root Cause Diagnosis of In-Production Failures** 344
Baris Kasikci, Benjamin Schubert (EPFL), Cristiano Pereira, Gilles Pokam (Intel Corporation), George Candea (EPFL)
- Cross-checking Semantic Correctness: The Case of Finding File System Bugs**..... 361
Changwoo Min, Sanidhya Kashyap, Byoungyoung Lee, Chengyu Song, Taesoo Kim (Georgia Institute of Technology)
- Pivot Tracing: Dynamic Causal Monitoring for Distributed Systems** 378
Jonathan Mace, Ryan Roelke, Rodrigo Fonseca (Brown University)

Big Data

Chair: Terence Kelly

- Interruptible Tasks: Treating Memory Pressure As Interrupts for Highly Scalable Data-Parallel Programs** 394
Lu Fang, Khanh Nguyen, Guoqing Xu, Brian Demsky (University of California, Irvine), Shan Lu (University of Chicago)
- Chaos: Scale-out Graph Processing from Secondary Storage** 410
Amitabha Roy (Intel), Laurent Bindschaedler, Jasmina Malicevic, Willy Zwaenepoel (EPFL)
- Arabesque: A System for Distributed Graph Mining** 425
Carlos H. C. Teixeira, Alexandre J. Fonseca, Marco Serafini, Georgos Siganos, Mohammed J. Zaki, Ashraf Aboulnaga (Qatar Computing Research Institute)

Storage Systems

Chair: Andrew Baumann

- How to Get More Value From Your File System Directory Cache** 441
Chia-Che Tsai, Yang Zhan, Jayashree Reddy, Yizheng Jiao, Tao Zhang, Donald E. Porter (Stony Brook University)
- Opportunistic Storage Maintenance** 457
George Amvrosiadis, Angela Demke Brown, Ashvin Goel (University of Toronto)
- Split-Level I/O Scheduling** 474
Suli Yang, Tyler Harter, Nishant Agrawal, Salini Selvaraj Kowsalya, Anand Krishnamurthy, Samer Al-Kiswany (University of Wisconsin-Madison), Rini T. Kaushik (IBM Research - Almaden), Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau (University of Wisconsin-Madison)

Sponsors



Supporters



Organizers

General Chair:	Ethan Miller, UC Santa Cruz
Program Chair:	Steven Hand, Google
Treasurer:	Jacob Lorch, MSR Redmond
Local Arrangements:	Kevin Greenan, UC Santa Cruz
Sponsorship Chair:	Edouard Bugnion, EPFL
Workshop Chair:	Peter Desnoyers, Northeastern
Scholarship Chair:	Nitin Agrawal, NEC
Poster Chair:	Daniel Peek, Facebook
Publications:	Andrew Birrell
Program Committee:	Marcos Aguilera, VMware Gustavo Alonso, ETH Zurich Andrew Baumann, MSR Redmond Herbert Bos, Vrije Universiteit Eric Brewer, UC Berkeley, Google Miguel Castro, MSR Cambridge Luis Ceze, U. Washington Haibo Chen, Shanghai Jiao Tong Landon Cox, Duke Angela Demke Brown, U. Toronto Ulfar Erlingsson, Google Alexandra Fedorova, Simon Fraser Bryan Ford, Yale Steven Hand, Google Tim Harris, Oracle Labs Michael Kaminsky, Intel Labs Brad Karp, University College Terence Kelly, HP Labs Eddie Kohler, Harvard James Larus, EPFL Shan Lu, U. Chicago Petros Maniatis, Google Jeff Mogul, Google Jason Nieh, Columbia Steve Swanson, UC San Diego Mike Swift, U. Wisconsin Michael Walfish, NYU Hakim Weatherspoon, Cornell Emmett Witchel, UT Austin Nickolai Zeldovich, MIT