

USENIX ATC '18:
2018 USENIX Annual Technical Conference
July 11–13, 2018
Boston, MA, USA

Performance

Tributary: spot-dancing for elastic services with latency SLOs1
Aaron Harlap and Andrew Chung, *Carnegie Mellon University*; Alexey Tumanov, *UC Berkeley*;
Gregory R. Ganger and Phillip B. Gibbons, *Carnegie Mellon University*

**FastTrack: Foreground App-Aware I/O Management for Improving User Experience
of Android Smartphones**15
Sangwook Shane Hahn, *Seoul National University*; Sungjin Lee, *DGIST*; Inhyuk Yee, *AIBrain Asia*;
Donguk Ryu, *Samsung Electronics*; Jihong Kim, *Seoul National University*

Mainstream: Dynamic Stem-Sharing for Multi-Tenant Video Processing29
Angela H. Jiang, Daniel L.K. Wong, Christopher Canel, Lilia Tang, and Ishan Misra, *Carnegie Mellon
University*; Michael Kaminsky, Michael A. Kozuch, and Padmanabhan Pillai, *Intel Labs*; David G. Andersen
and Gregory R. Ganger, *Carnegie Mellon University*

VIDEOCHEF: Efficient Approximation for Streaming Video Processing Pipelines43
Ran Xu, Jinkyu Koo, Rakesh Kumar, and Peter Bai, *Purdue University*; Subrata Mitra, *Adobe Research*;
Sasa Misailovic, *University of Illinois Urbana-Champaign*; Saurabh Bagchi, *Purdue University*

Kernel

SOCK: Rapid Task Provisioning with Serverless-Optimized Containers57
Edward Oakes, Leon Yang, Dennis Zhou, and Kevin Houck, *University of Wisconsin-Madison*; Tyler Harter,
Microsoft, GSL; Andrea C. Arpaci-Dusseau and Remzi H. Arpaci-Dusseau, *University of Wisconsin-Madison*

DynaMix: Dynamic Mobile Device Integration for Efficient Cross-device Resource Sharing71
Dongju Chae, *POSTECH*; Joonsung Kim and Gwangmu Lee, *Seoul National University*; Hanjun Kim, *POSTECH*;
Kyung-Ah Chang and Hyogun Lee, *Samsung Electronics*; Jangwoo Kim, *Seoul National University*

The Battle of the Schedulers: FreeBSD ULE vs. Linux CFS85
Justinien Bouron, Sebastien Chevalley, Baptiste Lepers, and Willy Zwaenepoel, *EPFL*; Redha Gouicem,
Julia Lawall, Gilles Muller, and Julien Sopena, *Sorbonne University/Inria/LIP6*

The Design and Implementation of Hyperupcalls97
Nadav Amit and Michael Wei, *VMware Research*

Security 1

AIQL: Enabling Efficient Attack Investigation from System Monitoring Data113
Peng Gao, *Princeton University*; Xusheng Xiao, *Case Western Reserve University*; Zhichun Li and
Kangkook Jee, *NEC Laboratories America, Inc.*; Fengyuan Xu, *National Key Lab for Novel Software
Technology, Nanjing University*; Sanjeev R. Kulkarni and Prateek Mittal, *Princeton University*

Application Memory Isolation on Ultra-Low-Power MCUs127
Taylor Hardin, *Dartmouth College*; Ryan Scott, *Clemson University*; Patrick Proctor, *Dartmouth College*;
Josiah Hester, *Northwestern University*; Jacob Sorber, *Clemson University*; David Kotz, *Dartmouth College*

Peeking Behind the Curtains of Serverless Platforms133
Liang Wang, *UW-Madison*; Mengyuan Li and Yinqian Zhang, *The Ohio State University*; Thomas Ristenpart,
Cornell Tech; Michael Swift, *UW-Madison*

SOTERIA: Automated IoT Safety and Security Analysis147
Z. Berkay Celik, Patrick McDaniel, and Gang Tan, *The Pennsylvania State University*

Virtualization

- Scaling Guest OS Critical Sections with eCS**159
Sanidhya Kashyap, *Georgia Institute of Technology*; Changwoo Min, *Virginia Tech*; Taesoo Kim, *Georgia Institute of Technology*
- KylinX: A Dynamic Library Operating System for Simplified and Efficient Cloud Virtualization** 173
Yiming Zhang, *NiceX Lab, NUDT*; Jon Crowcroft, *University of Cambridge*; Dongsheng Li and Chengfen Zhang, *NUDT*; Huiba Li, *Alibaba*; Yaozheng Wang and Kai Yu, *NUDT*; Yongqiang Xiong, *Microsoft*; Guihai Chen, *SJTU*
- Virtualizing Energy Storage Management Using RAIBA**187
Tzi-cker Chiueh, Mao-Cheng Huang, Kai-Cheung Juang, Shih-Hao Liang, and Welkin Ling, *Industrial Technology Research Institute*
- CNTR: Lightweight OS Containers**199
Jörg Thalheim and Pramod Bhatotia, *University of Edinburgh*; Pedro Fonseca, *University of Washington*; Baris Kasikci, *University of Michigan*

Security 2

- Throwhammer: Rowhammer Attacks over the Network and Defenses**213
Andrei Tatar and Radhesh Krishnan Konoth, *Vrije Universiteit Amsterdam*; Elias Athanasopoulos, *University of Cyprus*; Cristiano Giuffrida, Herbert Bos, and Kaveh Razavi, *Vrije Universiteit Amsterdam*
- Varys: Protecting SGX Enclaves from Practical Side-Channel Attacks**227
Oleksii Oleksenko, Bohdan Trach, Robert Krahn, and André Martin, *TU Dresden*; Mark Silberstein, *Technion*; Christof Fetzer, *TU Dresden*
- Kernel-Supported Cost-Effective Audit Logging for Causality Tracking**241
Shiqing Ma, *Purdue University*; Juan Zhai, *Nanjing University*; Yonghwi Kwon, *Purdue University*; Kyu Hyung Lee, *University of Georgia*; Xiangyu Zhang, *Purdue University*; Gabriela Ciocarlie, Ashish Gehani, and Vinod Yegneswaran, *SRI International*; Dongyan Xu, *Purdue University*; Somesh Jha, *University of Wisconsin-Madison*
- EPTI: Efficient Defence against Meltdown Attack for Unpatched VMs**255
Zhichao Hua, Dong Du, Yubin Xia, Haibo Chen, and Binyu Zang, *Institute of Parallel and Distributed Systems, Shanghai Jiao Tong University*

Multicore

- Effectively Mitigating I/O Inactivity in vCPU Scheduling**267
Weiwei Jia, *The University of Hong Kong, New Jersey Institute of Technology*; Cheng Wang and Xusheng Chen, *The University of Hong Kong*; Jianchen Shan and Xiaowei Shang, *New Jersey Institute of Technology*; Heming Cui, *The University of Hong Kong*; Xiaoning Ding, *New Jersey Institute of Technology*; Luwei Cheng, *Facebook*; Francis C. M. Lau and Yuexuan Wang, *The University of Hong Kong*; Yuangang Wang, *Huawei*
- Placement of Virtual Containers on NUMA systems: A Practical and Comprehensive Model**281
Justin Funston, Maxime Lorrillere, and Alexandra Fedorova, *University of British Columbia*; Baptiste Lepers, *EPFL*; David Vengerov and Jean-Pierre Lozi, *Oracle Labs*; Vivien Quéma, *IMAG*
- Getting to the Root of Concurrent Binary Search Tree Performance**295
Maya Arbel-Raviv, *Technion*; Trevor Brown, *IST Austria*; Adam Morrison, *Tel Aviv University*
- TerseCades: Efficient Data Compression in Stream Processing**307
Gennady Pekhimenko, *University of Toronto*; Chuanxiong Guo, *Bytedance Inc.*; Myeongjae Jeon, *Microsoft Research*; Peng Huang, *Johns Hopkins University*; Lidong Zhou, *Microsoft Research*

(continued on next page)

Problem Determination

- Troubleshooting Transiently-Recurring Errors in Production Systems with Blame-Proportional Logging** 321
Liang Luo, *University of Washington*; Suman Nath, Lenin Ravindranath Sivalingam, and Madan Musuvathi, *Microsoft Research*; Luis Ceze, *University of Washington*
- NanoLog: A Nanosecond Scale Logging System**335
Stephen Yang, Seo Jin Park, and John Ousterhout, *Stanford University*
- Model Governance: Reducing the Anarchy of Production ML**351
Vinay Sridhar, Sriram Subramanian, Dulcardo Arteaga, Swaminathan Sundararaman, Drew Roselli, and Nisha Talagala, *ParallelM*

Consistency

- Fine-grained consistency for geo-replicated systems**359
Cheng Li, *University of Science and Technology of China*; Nuno Pregoica, *NOVA LINCS & FCT, Univ. NOVA de Lisboa*; Rodrigo Rodrigues, *INESC-ID & Instituto Superior Técnico, Universidade de Lisboa*
- Log-Free Concurrent Data Structures**373
Tudor David, *IBM Research, Zurich*; Aleksandar Dragojevic, *MSR Cambridge*; Rachid Guerraoui and Igor Zablotchi, *EPFL*
- Stable and Consistent Membership at Scale with Rapid**387
Lalith Suresh, Dahlia Malkhi, and Parikshit Gopalan, *VMware Research*; Ivan Porto Carreiro, *One Concern*; Zeeshan Lokhandwala, *VMware*

Big Data Faster

- On Smart Query Routing: For Distributed Graph Querying with Decoupled Storage**401
Arijit Khan, *Nanyang Technological University, Singapore*; Gustavo Segovia, *ETH Zurich, Switzerland*; Donald Kossmann, *Microsoft Research, Redmond, USA*
- Locality-Aware Software Throttling for Sparse Matrix Operation on GPUs**413
Yanhao Chen and Ari B. Hayes, *Rutgers University*; Chi Zhang, *University of Pittsburgh*; Timothy Salmon and Eddy Z. Zhang, *Rutgers University*
- Accelerating PageRank using Partition-Centric Processing**427
Kartik Lakhotia, *University of Southern California*; Rajgopal Kannan, *US Army Research Lab*; Viktor Prasanna, *University of Southern California*
- CGraph: A Correlations-aware Approach for Efficient Concurrent Iterative Graph Processing**441
Yu Zhang, Xiaofei Liao, Hai Jin, and Lin Gu, *Huazhong University of Science and Technology*; Ligang He, *University of Warwick*; Bingsheng He, *National University of Singapore*; Haikun Liu, *Huazhong University of Science and Technology*

Availability

- Don't share, Don't lock: Large-scale Software Connection Tracking with Krononat**453
Fabien André, Stéphane Gouache, Nicolas Le Scouarnec, and Antoine Monsifrot, *Technicolor*
- Accurate Timeout Detection Despite Arbitrary Processing Delays**467
Sixiang Ma and Yang Wang, *The Ohio State University*
- Improving Service Availability of Cloud Systems by Predicting Disk Error**481
Yong Xu and Kaixin Sui, *Microsoft Research, China*; Randolph Yao, *Microsoft Azure, USA*; Hongyu Zhang, *The University of Newcastle, Australia*; Qingwei Lin, *Microsoft Research, China*; Yingnong Dang, *Microsoft Azure, USA*; Peng Li, *Nankai University, China*; Keceng Jiang, Wenchi Zhang, and Jian-Guang Lou, *Microsoft Research, China*; Murali Chintalapati, *Microsoft Azure, USA*; Dongmei Zhang, *Microsoft Research, China*

RAFI: Risk-Aware Failure Identification to Improve the RAS in Erasure-coded Data Centers495
Juntao Fang, *Wuhan National Laboratory for Optoelectronics, Huazhong University of Sci. and Tech.*;
Shenggang Wan, *School of Computer Science and Technology, Huazhong University of Sci. and Tech.*;
Xubin He, *Department of Computer and Information Sciences, Temple University*

Big Data 1

Siphon: Expediting Inter-Datacenter Coflows in Wide-Area Data Analytics507
Shuhao Liu, Li Chen, and Baochun Li, *University of Toronto*

PerfIso: Performance Isolation for Commercial Latency-Sensitive Services519
Călin Iorgulescu, *EPFL*; Reza Azimi, *Brown University*; Youngjin Kwon, *U. Texas at Austin*; Sameh Elnikety,
Manoj Syamala, and Vivek Narasayya, *Microsoft Research*; Herodotos Herodotou, *Cyprus University of
Technology*; Paulo Tomita, Alex Chen, Jack Zhang, and Junhua Wang, *Microsoft Bing*

On the diversity of cluster workloads and its impact on research results.533
George Amvrosiadis, Jun Woo Park, Gregory R. Ganger, and Garth A. Gibson, *Carnegie Mellon University*;
Elisabeth Baseman and Nathan DeBardeleben, *Los Alamos National Laboratory*

SLAOrchestrator: Reducing the Cost of Performance SLAs for Cloud Data Analytics547
Jennifer Ortiz, Brendan Lee, and Magdalena Balazinska, *University of Washington*; Johannes Gehrke, *Microsoft*;
Joseph L. Hellerstein, *eScience Institute*

Analyzing Code

Spindle: Informed Memory Access Monitoring561
Haojie Wang, *Tsinghua University, Qatar Computing Research Institute*; Jidong Zhai, *Tsinghua University*;
Xiongchao Tang, *Tsinghua University, Qatar Computing Research Institute*; Bowen Yu, *Tsinghua University*;
Xiaosong Ma, *Qatar Computing Research Institute*; Wenguang Chen, *Tsinghua University*

Touchstone: Generating Enormous Query-Aware Test Databases575
Yuming Li and Rong Zhang, *East China Normal University*; Xiaoyan Yang and Zhenjie Zhang, *Singapore R&D,
Yitu Technology Ltd.*; Aoying Zhou, *East China Normal University*

DSAC: Effective Static Analysis of Sleep-in-Atomic-Context Bugs in Kernel Modules587
Jia-Ju Bai and Yu-Ping Wang, *Tsinghua University*; Julia Lawall, *Sorbonne Université/Inria/LIP6*; Shi-Min Hu,
Tsinghua University

Coccinelle: 10 Years of Automated Evolution in the Linux Kernel601
Julia Lawall and Gilles Muller, *Sorbonne University/Inria/LIP6*

Big Data 2

Albis: High-Performance File Format for Big Data Systems615
Animesh Trivedi, Patrick Stuedi, Jonas Pfefferle, Adrian Schuepbach, and Bernard Metzler, *IBM Research, Zurich*

Litz: Elastic Framework for High-Performance Distributed Machine Learning631
Aurick Qiao, *Petuum, Inc. and Carnegie Mellon University*; Abutalib Aghayev, *Carnegie Mellon University*;
Weiren Yu, *Petuum, Inc. and Beihang University*; Haoyang Chen and Qirong Ho, *Petuum, Inc.*;
Garth A. Gibson, *Carnegie Mellon University and Vector Institute*; Eric P. Xing, *Petuum, Inc. and Carnegie
Mellon University*

Putting the “Micro” Back in Microservice645
Sol Boucher, Anuj Kalia, and David G. Andersen, *Carnegie Mellon University*; Michael Kaminsky, *Intel Labs*

Fast and Concurrent RDF Queries using RDMA-assisted GPU Graph Exploration651
Siyuan Wang, Chang Lou, Rong Chen, and Haibo Chen, *Shanghai Jiao Tong University*

(continued on next page)

SSDs

- MDev-NVMe: A NVMe Storage Virtualization Solution with Mediated Pass-Through**665
Bo Peng, *Shanghai Jiao Tong University, Intel*; Haozhong Zhang, *Intel*; Jianguo Yao, *Shanghai Jiao Tong University*; Yaozu Dong, *Intel*; Yu Xu and Haibing Guan, *Shanghai Jiao Tong University*
- AutoSSD: an Autonomic SSD Architecture**677
Bryan S. Kim, *Seoul National University*; Hyun Suk Yang, *Hongik University*; Sang Lyul Min, *Seoul National University*
- Geriatric: Aging what you see and what you don't see. A file system aging approach for modern storage systems**691
Saurabh Kadekodi, Vaishnavh Nagarajan, and Gregory R. Ganger, *Carnegie Mellon University*;
Garth A. Gibson, *Carnegie Mellon University, Vector Institute*
- Can't We All Get Along? Redesigning Protection Storage for Modern Workloads**705
Yamini Allu, Fred Douglass, Mahesh Kamat, Ramya Prabhakar, Philip Shilane, and Rahul Ugale, *Dell EMC*

The Network

- STMS: Improving MPTCP Throughput Under Heterogeneous Networks**719
Hang Shi and Yong Cui, *Tsinghua University*; Xin Wang, *Stony Brook University*; Yuming Hu and Minglong Dai, *Tsinghua University*; Fanzhao Wang and Kai Zheng, *Huawei Technologies*
- Pantheon: the training ground for Internet congestion-control research**731
Francis Y. Yan, Jestin Ma, and Greg D. Hill, *Stanford University*; Deepti Raghavan, *Massachusetts Institute of Technology*; Riad S. Wahby, Philip Levis, and Keith Winstein, *Stanford University*
- ClickNF: a Modular Stack for Custom Network Functions**745
Massimo Gallo and Rafael Laufer, *Nokia Bell Labs*

Storage 1

- Selecta: Heterogeneous Cloud Storage Configuration for Data Analytics**759
Ana Klimovic, *Stanford University*; Heiner Litz, *UC Santa Cruz*; Christos Kozyrakis, *Stanford University*
- Remote regions: a simple abstraction for remote memory**775
Marcos K. Aguilera, Nadav Amit, Irina Calciu, Xavier Deguillard, Jayneel Gandhi, Stanko Novakovic, Arun Ramanathan, Pratap Subrahmanyam, Lalith Suresh, Kiran Tati, Rajesh Venkatasubramanian, and Michael Wei, *VMware*
- Understanding Ephemeral Storage for Serverless Analytics**789
Ana Klimovic, Yawen Wang, and Christos Kozyrakis, *Stanford University*; Patrick Stuedi, Jonas Pfefferle, and Animesh Trivedi, *IBM Research*

Transactions

- Solar: Towards a Shared-Everything Database on Distributed Log-Structured Storage**795
Tao Zhu, *East China Normal University*; Zhuoyue Zhao and Feifei Li, *University of Utah*; Weining Qian and Aoying Zhou, *East China Normal University*; Dong Xie and Ryan Stutsman, *University of Utah*; Haining Li, *Bank of Communications*; Huiqi Hu, *East China Normal University*; *Bank of Communications*
- Toward Coordination-free and Reconfigurable Mixed Concurrency Control**809
Dixin Tang and Aaron J. Elmore, *University of Chicago*
- Scaling Hardware Accelerated Network Monitoring to Concurrent and Dynamic Queries With *Flow**823
John Sonchack, *University of Pennsylvania*; Oliver Michel, *University of Colorado Boulder*; Adam J. Aviv, *United States Naval Academy*; Eric Keller, *University of Colorado Boulder*; Jonathan M. Smith, *University of Pennsylvania*
- Applying Hardware Transactional Memory for Concurrency-Bug Failure Recovery in Production Runs** ... 837
Yuxi Chen, Shu Wang, and Shan Lu, *University of Chicago*; Karthikeyan Sankaralingam, *University of Wisconsin – Madison*

Storage 2

Tailwind: Fast and Atomic RDMA-based Replication851
Yacine Taleb, *Univ Rennes, Inria, CNRS, IRISA*; Ryan Stutsman, *University of Utah*; Gabriel Antoniu, *Univ Rennes, Inria, CNRS, IRISA*; Toni Cortes, *BSC, UPC*

On Fault Tolerance, Locality, and Optimality in Locally Repairable Codes865
Oleg Kolosov, *School of Electrical Engineering, Tel Aviv University*; Gala Yadgar, *Computer Science Department, Technion, and School of Electrical Engineering, Tel Aviv University*; Matan Liram, *Computer Science Department, Technion*; Itzhak Tamo, *School of Electrical Engineering, Tel Aviv University*; Alexander Barg, *Department of ECE/ISR, University of Maryland*

TxFs: Leveraging File-System Crash Consistency to Provide ACID Transactions879
Yige Hu, Zhiting Zhu, Ian Neal, Youngjin Kwon, and Tianyu Cheng, *The University of Texas at Austin*; Vijay Chidambaram, *The University of Texas at Austin and VMware Research*; Emmett Witchel, *The University of Texas at Austin*

Towards Better Understanding of Black-box Auto-Tuning: A Comparative Analysis for Storage Systems ... 893
Zhen Cao, *Stony Brook University*; Vasily Tarasov, *IBM Research–Almaden*; Sachin Tiwari and Erez Zadok, *Stony Brook University*

Data Center/Machine Learning

HeavyKeeper: An Accurate Algorithm for Finding Top-*k* Elephant Flows909
Junzhi Gong, Tong Yang, Haowei Zhang, and Hao Li, *Peking University*; Steve Uhlig, *Queen Mary, University of London*; Shigang Chen, *University of Florida*; Lorna Uden, *Staffordshire University*; Xiaoming Li, *Peking University*

SAND: Towards High-Performance Serverless Computing923
Istemi Ekin Akkus, Ruichuan Chen, Ivica Rimac, Manuel Stein, Klaus Satzke, Andre Beck, Paarijaat Aditya, and Volker Hilt, *Nokia Bell Labs*

Cavs: An Efficient Runtime System for Dynamic Neural Networks937
Shizhen Xu, *Carnegie Mellon University, Tsinghua University*; Hao Zhang, Graham Neubig, and Wei Dai, *Carnegie Mellon University, Petuum Inc.*; Jin Kyu Kim, *Carnegie Mellon University*; Zhijie Deng, *Tsinghua University*; Qirong Ho, *Petuum Inc.*; Guangwen Yang, *Tsinghua University*; Eric P. Xing, *Petuum Inc.*

DeepCPU: Serving RNN-based Deep Learning Models 10x Faster951
Minjia Zhang, Samyam Rajbhandari, Wenhan Wang, and Yuxiong He, *Microsoft AI and Research*

Key/Value Storage

Closing the Performance Gap Between Volatile and Persistent Key-Value Stores Using Cross-Referencing Logs967
Yihe Huang, *Harvard University*; Matej Pavlovic, *EPFL*; Virendra Marathe, Margo Seltzer, Tim Harris, and Steve Blyan, *Oracle Labs*

Metis: Robustly Tuning Tail Latencies of Cloud Systems981
Zhao Lucis Li, *USTC*; Chieh-Jan Mike Liang, *Microsoft Research*; Wenjia He, *USTC*; Lianjie Zhu, Wenjun Dai, and Jin Jiang, *Microsoft Bing Ads*; Guangzhong Sun, *USTC*

Redesigning LSMs for Nonvolatile Memory with NoveLSM993
Sudarsun Kannan, *University of Wisconsin-Madison*; Nitish Bhat and Ada Gavrilovska, *Georgia Tech*; Andrea Arpaci-Dusseau and Remzi Arpaci-Dusseau, *University of Wisconsin-Madison*

HashKV: Enabling Efficient Updates in KV Storage via Hashing1007
Helen H. W. Chan, *The Chinese University of Hong Kong*; Yongkun Li, *University of Science and Technology of China*; Patrick P. C. Lee, *The Chinese University of Hong Kong*; Yinlong Xu, *University of Science and Technology of China*