

16th USENIX Symposium on Operating Systems
Design and Implementation (OSDI '22)

July 11–13, 2022

Carlsbad, CA, USA

Monday, July 11

Distributed Storage and Far Memory

Owl: Scale and Flexibility in Distribution of Hot Content 1
Jason Flinn, Xianzheng Dou, Arushi Aggarwal, Alex Boyko, Francois Richard, Eric Sun, Wendy Tobagus, Nick Wolchko, and Fang Zhou, *Meta*

BlockFlex: Enabling Storage Harvesting with Software-Defined Flash in Modern Cloud Platforms 17
Benjamin Reidys and Jinghan Sun, *University of Illinois at Urbana-Champaign*; Anirudh Badam and Shadi Noghahi, *Microsoft Research*; Jian Huang, *University of Illinois at Urbana-Champaign*

MemLiner: Lining up Tracing and Application for a Far-Memory-Friendly Runtime 35
Chenxi Wang, Haoran Ma, Shi Liu, Yifan Qiao, Jonathan Eyolfson, and Christian Navasca, *UCLA*; Shan Lu, *University of Chicago*; Guoqing Harry Xu, *UCLA*

Carbink: Fault-Tolerant Far Memory 55
Yang Zhou, *Harvard University*; Hassan M. G. Wassel, *Google*; Sihang Liu, *University of Virginia*; Jiaqi Gao and James Mickens, *Harvard University*; Minlan Yu, *Harvard University and Google*; Chris Kennelly, Paul Turner, and David E. Culler, *Google*; Henry M. Levy, *University of Washington and Google*; Amin Vahdat, *Google*

Bugs

Metastable Failures in the Wild 73
Lexiang Huang, *The Pennsylvania State University and Twitter*; Matthew Magnusson and Abishek Bangalore Muralikrishna, *University of New Hampshire*; Salman Estyak, *The Pennsylvania State University*; Rebecca Isaacs, *Twitter*; Abutalib Aghayev and Timothy Zhu, *The Pennsylvania State University*; Aleksey Charapko, *University of New Hampshire*

Demystifying and Checking Silent Semantic Violations in Large Distributed Systems 91
Chang Lou, Yuzhuo Jing, and Peng Huang, *Johns Hopkins University*

RESIN: A Holistic Service for Dealing with Memory Leaks in Production Cloud Infrastructure 109
Chang Lou, *Johns Hopkins University*; Cong Chen, *Microsoft Azure*; Peng Huang, *Johns Hopkins University*; Yingnong Dang, *Microsoft Azure*; Si Qin, *Microsoft Research*; Xinsheng Yang, *Meta*; Xukun Li, *Microsoft Azure*; Qingwei Lin, *Microsoft Research*; Murali Chintalapati, *Microsoft Azure*

Cancellation in Systems: An Empirical Study of Task Cancellation Patterns and Failures 127
Utsav Sethi and Haochen Pan, *University of Chicago*; Shan Lu, *University of Chicago and Microsoft*; Madanlal Musuvathi and Suman Nath, *Microsoft Research*

Automatic Reliability Testing For Cluster Management Controllers 143
Xudong Sun, Wenqing Luo, and Jiawei Tyler Gu, *University of Illinois at Urbana-Champaign*; Aishwarya Ganesan, Ramnatthan Alagappan, Michael Gasch, and Lalith Suresh, *VMware*; Tianyin Xu, *University of Illinois at Urbana-Champaign*

Persistent Memory

ListDB: Union of Write-Ahead Logs and Persistent SkipLists for Incremental Checkpointing on Persistent Memory 161
Wonbae Kim, *UNIST*; Chanyeol Park, *Sungkyunkwan University and Naver*; Dongui Kim, *Sungkyunkwan University and Line*; Hyeongjun Park, *Sungkyunkwan University*; Young-ri Choi, *UNIST*; Alan Sussman, *University of Maryland, College Park*; Beomseok Nam, *Sungkyunkwan University*

ODINFS: Scaling PM Performance with Opportunistic Delegation 179
Diyu Zhou, Yuchen Qian, Vishal Gupta, and Zhifei Yang, *EPFL*; Changwoo Min, *Virginia Tech*; Sanidhya Kashyap, *EPFL*

DURINN: Adversarial Memory and Thread Interleaving for Detecting Durable Linearizability Bugs 195
Xinwei Fu, *Virginia Tech*; Dongyoon Lee, *Stony Brook University*; Changwoo Min, *Virginia Tech*

Machine Learning 1

- SparTA: Deep-Learning Model Sparsity via Tensor-with-Sparsity-Attribute** 213
Ningxin Zheng, *Microsoft Research*; Bin Lin, *Microsoft Research and Tsinghua University*; Quanlu Zhang, Lingxiao Ma, Yuqing Yang, Fan Yang, Yang Wang, Mao Yang, and Lidong Zhou, *Microsoft Research*
- ROLLER: Fast and Efficient Tensor Compilation for Deep Learning** 233
Hongyu Zhu, *University of Toronto and Microsoft Research*; Ruofan Wu, *Renmin University of China and Microsoft Research*; Yijia Diao, *Shanghai Jiao Tong University and Microsoft Research*; Shanbin Ke, *UCSD and Microsoft Research*; Haoyu Li, *Columbia University and Microsoft Research*; Chen Zhang, *Tsinghua University and Microsoft Research*; Jilong Xue, Lingxiao Ma, Yuqing Xia, Wei Cui, Fan Yang, Mao Yang, and Lidong Zhou, *Microsoft Research*; Asaf Cidon, *Columbia University*; Gennady Pekhimenko, *University of Toronto*
- Walle: An End-to-End, General-Purpose, and Large-Scale Production System for Device-Cloud Collaborative Machine Learning** 249
Chengfei Lv, *Zhejiang University & Alibaba Group*; Chaoyue Niu, *Shanghai Jiao Tong University & Alibaba Group*; Renjie Gu, Xiaotang Jiang, Zhaode Wang, Bin Liu, Ziqi Wu, Qiulin Yao, Congyu Huang, Panos Huang, Tao Hudyang, Hui Shu, Jinde Song, Bin Zou, Peng Lan, and Guohuan Xu, *Alibaba Group*; Fei Wu, *Zhejiang University*; Shaojie Tang, *University of Texas at Dallas*; Fan Wu and Guihai Chen, *Shanghai Jiao Tong University*
- Unity: Accelerating DNN Training Through Joint Optimization of Algebraic Transformations and Parallelization** 267
Colin Unger, *Stanford University*; Zhihao Jia, *Carnegie Mellon University and Meta*; Wei Wu, *Los Alamos National Laboratory and NVIDIA*; Sina Lin, *Microsoft*; Mandeep Baines and Carlos Efrain Quintero Narvaez, *Meta*; Vinay Ramakrishnaiah, Nirmal Prajapati, Pat McCormick, and Jamaludin Mohd-Yusof, *Los Alamos National Laboratory*; Xi Luo, *SLAC National Accelerator Laboratory*; Dheevatsa Mudigere, Jongsoo Park, and Misha Smelyanskiy, *Meta*; Alex Aiken, *Stanford University*

Tuesday, July 12

Potpourri

- Trinity: High-Performance Mobile Emulation through Graphics Projection** 285
Di Gao, Hao Lin, Zhenhua Li, Chengen Huang, and Yunhao Liu, *Tsinghua University*; Feng Qian, *University of Minnesota*; Liangyi Gong, *CNIC, CAS*; Tianyin Xu, *UIUC*
- ORION and the Three Rights: Sizing, Bundling, and Prewarming for Serverless DAGs** 303
Ashraf Mahgoub and Edgardo Barsallo Yi, *Purdue University*; Karthick Shankar, *Carnegie Mellon University*; Sameh Elnikety, *Microsoft Research*; Somali Chaterji and Saurabh Bagchi, *Purdue University*
- Occualizer: Optimistic Concurrent Search Trees From Sequential Code** 321
Tomer Shanny and Adam Morrison, *Tel Aviv University*
- Immortal Threads: Multithreaded Event-driven Intermittent Computing on Ultra-Low-Power Microcontrollers** . 339
Eren Yıldız, *Ege University*; Lijun Chen and Kasim Sinan Yıldırım, *University of Trento*
- Debugging the OmniTable Way** 357
Andrew Quinn, *UC Santa Cruz*; Jason Flinn, *Meta*; Michael Cafarella, *MIT*; Baris Kasikci, *University of Michigan*

Storage

- XRP: In-Kernel Storage Functions with eBPF** 375
Yuhong Zhong, Haoyu Li, Yu Jian Wu, Ioannis Zarkadas, Jeffrey Tao, Evan Mesterhazy, Michael Makris, and Junfeng Yang, *Columbia University*; Amy Tai, *Google*; Ryan Stutsman, *University of Utah*; Asaf Cidon, *Columbia University*
- TriCache: A User-Transparent Block Cache Enabling High-Performance Out-of-Core Processing with In-Memory Programs** 395
Guanyu Feng and Huanqi Cao, *Tsinghua University*; Xiaowei Zhu, *Ant Group*; Bowen Yu, Yuanwei Wang, Zixuan Ma, Shengqi Chen, and Wenguang Chen, *Tsinghua University*
- Tiger: Disk-Adaptive Redundancy without Placement Restrictions** 413
Saurabh Kadekodi, *Google*; Francisco Maturana and Sanjith Athlur, *Carnegie Mellon University*; Arif Merchant, *Google*; K. V. Rashmi and Gregory R. Ganger, *Carnegie Mellon University*
- zIO: Accelerating IO-Intensive Applications with Transparent Zero-Copy IO** 431
Timothy Stamler, Deukyeon Hwang, and Amanda Raybuck, *UT Austin*; Wei Zhang, *Microsoft*; Simon Peter, *University of Washington*

Formal Verification

- Verifying the DaisyNFS concurrent and crash-safe file system with sequential reasoning** 447
Tej Chajed, *MIT CSAIL*; Joseph Tassarotti, *Boston College*; Mark Theng, M. Frans Kaashoek, and Nickolai Zeldovich, *MIT CSAIL*
- Design and Verification of the Arm Confidential Compute Architecture** 465
Xupeng Li and Xuheng Li, *Columbia University*; Christoffer Dall, *Arm Ltd*; Ronghui Gu and Jason Nieh, *Columbia University*; Yousuf Sait and Gareth Stockwell, *Arm Ltd*
- DuoAI: Fast, Automated Inference of Inductive Invariants for Verifying Distributed Protocols** 485
Jianan Yao, Runzhou Tao, Ronghui Gu, and Jason Nieh, *Columbia University*
- Verifying Hardware Security Modules with Information-Preserving Refinement** 503
Anish Athalye, M. Frans Kaashoek, and Nickolai Zeldovich, *MIT CSAIL*

Machine Learning 2

- ORCA: A Distributed Serving System for Transformer-Based Generative Models** 521
Gyeong-In Yu and Joo Seong Jeong, *Seoul National University*; Geon-Woo Kim, *FriendliAI and Seoul National University*; Soojeong Kim, *FriendliAI*; Byung-Gon Chun, *FriendliAI and Seoul National University*
- Microsecond-scale Preemption for Concurrent GPU-accelerated DNN Inferences** 539
Mingcong Han, *Institute of Parallel and Distributed Systems, SEIEE, Shanghai Jiao Tong University; Shanghai AI Laboratory*; Hanze Zhang, *Institute of Parallel and Distributed Systems, SEIEE, Shanghai Jiao Tong University; MoE Key Lab of Artificial Intelligence, AI Institute, Shanghai Jiao Tong University, China*; Rong Chen, *Institute of Parallel and Distributed Systems, SEIEE, Shanghai Jiao Tong University; Shanghai AI Laboratory*; Haibo Chen, *Institute of Parallel and Distributed Systems, SEIEE, Shanghai Jiao Tong University; Engineering Research Center for Domain-specific Operating Systems, Ministry of Education, China*
- Alpa: Automating Inter- and Intra-Operator Parallelism for Distributed Deep Learning** 559
Lianmin Zheng, Zhuohan Li, and Hao Zhang, *UC Berkeley*; Yonghao Zhuang, *Shanghai Jiao Tong University*; Zhifeng Chen and Yanping Huang, *Google*; Yida Wang, *Amazon Web Services*; Yuanzhong Xu, *Google*; Danyang Zhuo, *Duke University*; Eric P. Xing, *MBZUAI and Carnegie Mellon University*; Joseph E. Gonzalez and Ion Stoica, *UC Berkeley*
- Looking Beyond GPUs for DNN Scheduling on Multi-Tenant Clusters** 579
Jayashree Mohan, Amar Phanishayee, and Janardhan Kulkarni, *Microsoft Research*; Vijay Chidambaram, *The University of Texas at Austin and VMware Research*

Wednesday, July 13

Isolation and OS Services

- CAP-VMs: Capability-Based Isolation and Sharing in the Cloud** 597
Vasily A. Sartakov and Lluís Vilanova, *Imperial College London*; David Eyers, *University of Otago*; Takahiro Shinagawa, *The University of Tokyo*; Peter Pietzuch, *Imperial College London*
- KSplit: Automating Device Driver Isolation** 613
Yongzhe Huang, *Penn State University*; Vikram Narayanan and David Detweiler, *University of California, Irvine*; Kaiming Huang, Gang Tan, and Trent Jaeger, *Penn State University*; Anton Burtsev, *University of California, Irvine, and University of Utah*
- Operating System Support for Safe and Efficient Auxiliary Execution** 633
Yuzhuo Jing and Peng Huang, *Johns Hopkins University*
- From Dynamic Loading to Extensible Transformation: An Infrastructure for Dynamic Library Transformation...** 649
Yuxin Ren, Kang Zhou, Jianhai Luan, Yunfeng Ye, Shiyuan Hu, Xu Wu, Wenqin Zheng, Wenfeng Zhang, and Xinwei Hu, *Poincare lab, Huawei Technologies Co., Ltd, China*
- Application-Informed Kernel Synchronization Primitives** 667
Sujin Park, *Georgia Tech*; Diyu Zhou and Yuchen Qian, *EPFL*; Irina Calciu, *Graft*; Taesoo Kim, *Georgia Tech*; Sanidhya Kashyap, *EPFL*

Security and Private Messaging

- BlackBox: A Container Security Monitor for Protecting Containers on Untrusted Operating Systems** 683
Alexander Van't Hof and Jason Nieh, *Columbia University*
- Blockaid: Data Access Policy Enforcement for Web Applications**..... 701
Wen Zhang, *UC Berkeley*; Eric Sheng, *Yugabyte*; Michael Chang, *UC Berkeley*; Aurojit Panda, *NYU*; Mooly Sagiv, *Tel Aviv University*; Scott Shenker, *UC Berkeley/ICSI*
- SHORTSTACK: Distributed, Fault-tolerant, Oblivious Data Access**..... 719
Midhul Vuppalapati and Kushal Babel, *Cornell University*; Anurag Khandelwal, *Yale University*; Rachit Agarwal, *Cornell University*
- Groove: Flexible Metadata-Private Messaging** 735
Ludovic Barman, *EPFL*; Moshe Kol, *Hebrew University of Jerusalem*; David Lazar, *EPFL*; Yossi Gilad, *Hebrew University of Jerusalem*; Nikolai Zeldovich, *MIT CSAIL*

Managed Languages

- UPGRADVISOR: Early Adopting Dependency Updates Using Hybrid Program Analysis and Hardware Tracing**751
Yaniv David, *Columbia University*; Xudong Sun, *Nanjing University*; Raphael J. Sofaer, *Columbia University*; Aditya Senthilnathan, *IIT, Delhi*; Junfeng Yang, *Columbia University*; Zhiqiang Zuo, *Nanjing University*; Guoqing Harry Xu, *UCLA*; Jason Nieh and Ronghui Gu, *Columbia University*
- Practically Correct, Just-in-Time Shell Script Parallelization**..... 769
Konstantinos Kallas, *University of Pennsylvania*; Tammam Mustafa, *MIT CSAIL*; Jan Bielak, *XIV Staszic High School*; Dimitris Karnikis, *Aarno Labs*; Thurston H.Y. Dang, *MIT CSAIL*; Michael Greenberg, *Stevens Institute of Technology*; Nikos Vasilakis, *MIT CSAIL*
- Hubble: Performance Debugging with In-Production, Just-In-Time Method Tracing on Android** 787
Yu Luo and Kirk Rodrigues, *University of Toronto*; Cuiqin Li, Feng Zhang, Lijin Jiang, and Bing Xia, *Huawei Technologies Co., Ltd.*; David Lion and Ding Yuan, *University of Toronto*
- Jawa: Web Archival in the Era of JavaScript** 805
Ayush Goel and Jingyuan Zhu, *University of Michigan*; Ravi Netravali, *Princeton University*; Harsha V. Madhyastha, *University of Michigan*

Recommenders and Pattern Mining

- Ekko: A Large-Scale Deep Learning Recommender System with Low-Latency Model Update** 821
Chijun Sima, *Tencent*; Yao Fu and Man-Kit Sit, *The University of Edinburgh*; Liyi Guo, Xuri Gong, Feng Lin, Junyu Wu, Yongsheng Li, and Haidong Rong, *Tencent*; Pierre-Louis Aublin, *IJJ research laboratory*; Luo Mai, *The University of Edinburgh*
- FAERY: An FPGA-accelerated Embedding-based Retrieval System**..... 841
Chaoliang Zeng, *Hong Kong University of Science and Technology*; Layong Luo, Qingsong Ning, Yaodong Han, and Yuhang Jiang, *ByteDance*; Ding Tang, Zilong Wang, and Kai Chen, *Hong Kong University of Science and Technology*; Chuanxiong Guo, *ByteDance*
- Efficient and Scalable Graph Pattern Mining on GPUs**..... 857
Xuhao Chen and Arvind, *MIT CSAIL*