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 Noghabi, *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

**Bugs**

**Metastable Failures in the Wild** .......................................................... 73

**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

**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, Ramnath 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*

**OMNIS: 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*
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 CPUs for DNN Scheduling on Multi-Tenant Clusters ..................... 579
Jayashree Mohan, Amar Phanishayee, and Janardhan Kulkarni, Microsoft Research; Vijay Chidambaran, 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

Alexander Van’t Hof and Jason Nieh, Columbia University

Blockaid: Data Access Policy Enforcement for Web Applications

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

Midhul Vuppalapati and Kushal Babel, Cornell University; Anurag Khandelwal, Yale University; Rachit Agarwal, Cornell University

Groove: Flexible Metadata-Private Messaging

Ludovic Barman, EPFL; Moshe Kol, Hebrew University of Jerusalem; David Lazar, EPFL; Yossi Gilad, Hebrew University of Jerusalem; Nickolai Zeldovich, MIT CSAIL

Managed Languages

UpgAdvisor: Early Adopting Dependency Updates Using Hybrid Program Analysis and Hardware Tracing

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

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

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

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

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, IIJ research laboratory; Luo Mai, The University of Edinburgh

Maery: An FPGA-accelerated Embedding-based Retrieval System

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

Xuhao Chen and Arvind, MIT CSAIL