Monday, July 10

Make Your Bits Go Faster

Ship your Critical Section, Not Your Data: Enabling Transparent Delegation with TCLOCKS
Vishal Gupta, EPFL; Kumar Kartikeya Dwivedi, SRMIST; Yugesh Kothari, Yueyang Pan, Diyu Zhou, and Sanidhya Kashyap, EPFL

RON: One-Way Circular Shortest Routing to Achieve Efficient and Bounded-waiting Spinlocks
Shiwu Lo, Han-Ting Lin, Yao-Hung Hsieh, and Chao-Ting Lin, National Chung Cheng University; Yu-Hsueh Fang, National Cheng Kung University; Ching-Shen Lin, National Chung Cheng University; Ching-Chun (Jim) Huang, National Cheng Kung University; Kam Yiu Lam, City University of Hong Kong; Yuan-Hao Chang, Academia Sinica, Taiwan

Userspace Bypass: Accelerating Syscall-intensive Applications
Zhe Zhou, Yanxiang Bi, Junpeng Wan, and Yangfan Zhou, Fudan University; Zhou Li, University of California, Irvine

Triangulating Python Performance Issues with SCALENE
Emery D. Berger, Sam Stern, and Juan Altmayer Pizzorno, University of Massachusetts Amherst

Relational Debugging — Pinpointing Root Causes of Performance Problems
Xiang (Jenny) Ren, Sitao Wang, Zhuqi Jin, David Lion, and Adrian Chiu, University of Toronto; Tianyin Xu, University of Illinois at Urbana-Champaign; Ding Yuan, University of Toronto

Secure Your Bits I

Accountable authentication with privacy protection: The Larch system for universal login
Emma Dauterman, UC Berkeley; Danny Lin, Woodinville High School; Henry Corrigan-Gibbs, MIT; David Mazières, Stanford University

K9db: Privacy-Compliant Storage For Web Applications By Construction
Kinan Dak Albab, Ishan Sharma, Justus Adam, Benjamin Kilimnik, Aaron Jeyaraj, Raj Paul, Artem Agvanian, Leonhard Spiegelberg, and Malte Schwarzkopf, Brown University

Encrypted Databases Made Secure Yet Maintainable
Mingyu Li, Shanghai Jiao Tong University; Shanghai AI Laboratory; Engineering Research Center for Domain-specific Operating Systems, Ministry of Education, China; Xinyang Zhao and Le Chen, Shanghai Jiao Tong University; Engineering Research Center for Domain-specific Operating Systems, Ministry of Education, China; Cheng Tan, Northeastern University; Huorong Li and Sheng Wang, Alibaba Group; Zeyu Mi, Shanghai Jiao Tong University; Engineering Research Center for Domain-specific Operating Systems, Ministry of Education, China; Yubin Xia, Shanghai Jiao Tong University; Shanghai AI Laboratory; Engineering Research Center for Domain-specific Operating Systems, Ministry of Education, China; Feifei Li, Alibaba Group; Haibo Chen, Shanghai Jiao Tong University; Engineering Research Center for Domain-specific Operating Systems, Ministry of Education, China

LVMT: An Efficient Authenticated Storage for Blockchain
Chenxing Li, Shanghai Tree-Graph Blockchain Research Institute; Sidi Mohamed Beillahi, University of Toronto; Guang Yang and Ming Wu, Shanghai Tree-Graph Blockchain Research Institute; Wei Xu, Tsinghua University; Fan Long, University of Toronto

Honeycomb: Secure and Efficient GPU Executions via Static Validation
Haohui Mai, PrivacyCore Inc.; Jiacheng Zhao, SKLP, Institute of Computing Technology, CAS; Zhongguancun Laboratory; and UCAS; Hongren Zheng, IIS, Tsinghua University; Yiyang Zhao, SKLP, Institute of Computing Technology, CAS; and UCAS; Zibin Liu, BUPT; Mingyu Gao, IIS, Tsinghua University; Cong Wang, IDEA Shenzhen; Huimin Cui, SKLP, Institute of Computing Technology, CAS; and UCAS; Xiaobing Feng, SKLP, Institute of Computing Technology, CAS; Zhongguancun Laboratory; and UCAS; Christos Kozyrakis, PrivacyCore Inc. and Stanford
Secure Your Bits II
An Extensible Orchestration and Protection Framework for Confidential Cloud Computing .................173
Adil Ahmad and Alex Schultz, Arizona State University; Byoungyoung Lee, Seoul National University; Pedro Fonseca, Purdue University

Nimble: Rollback Protection for Confidential Cloud Services .......................................................... 193
Sebastian Angel, Microsoft Research; Aditya Basu, Penn State University; Weidong Cui, Microsoft Research; Trent Jaeger, Penn State University; Stella Lau, MIT CSAIL; Srinath Setty, Microsoft Research; Sudheesh Singanamalla, University of Washington

Kerveros: Efficient and Scalable Cloud Admission Control .......................................................... 209
Sultan Mahmud Sajal, Microsoft Research and Pennsylvania State University; Luke Marshall and Beibin Li, Microsoft Research; Shandan Zhou and Abhisek Pan, Microsoft Azure; Konstantina Mellou and Deepak Narayanan, Microsoft Research; Timothy Zhu, Pennsylvania State University; David Dion and Thomas Moscibroda, Microsoft Azure; Ishai Menache, Microsoft Research

Security and Performance in the Delegated User-level Virtualization ........................................ 227
Jiahaoo Chen, Institute of Parallel and Distributed Systems, SEIEE, Shanghai Jiao Tong University; Engineering Research Center for Domain-specific Operating Systems, Ministry of Education, China; Dingji Li, Institute of Parallel and Distributed Systems, SEIEE, Shanghai Jiao Tong University; Engineering Research Center for Domain-specific Operating Systems, Ministry of Education, China; MoE Key Lab of Artificial Intelligence, AI Institute, Shanghai Jiao Tong University; Zeyu Mi, Yuxuan Liu, and Binyu Zang, Institute of Parallel and Distributed Systems, SEIEE, Shanghai Jiao Tong University; Engineering Research Center for Domain-specific Operating Systems, Ministry of Education, China; Haibing Guan, Shanghai Key Laboratory of Scalable Computing and Systems, Shanghai Jiao Tong University; 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

Core slicing: closing the gap between leaky confidential VMs and bare-metal cloud ...................... 247
Ziqiao Zhou, Microsoft Research; Yizhou Shan, University of California, San Diego; Weidong Cui, Xinyang Ge, Marcus Peinado, and Andrew Baumann, Microsoft Research

Tuesday, July 11
Expanding, Hardening, and Deploying Your Bits

ExoFlow: A Universal Workflow System for Exactly-Once DAGs .................................................. 269
Siyuan Zhuang, UC Berkeley; Stephanie Wang, UC Berkeley and Anycale; Eric Liang and Yi Cheng, Anycale; Ion Stoica, UC Berkeley

Hyrax: Fail-in-Place Server Operation in Cloud Platforms ............................................................. 287
Jialun Lyu, Microsoft Azure and University of Toronto; Marisa You, Celine Irvene, Mark Jung, Tyler Narmore, Jacob Shapiro, Luke Marshall, and Savvasachi Samal, Microsoft Azure; Ioannis Manousakis and Lisa Hsu, Formerly of Microsoft Azure; Preetha Subbarayalu, Ashish Raniwala, Brijesh Warrier, and Ricardo Bianchini, Microsoft Azure; Bianca Schroeder, University of Toronto; Daniel S. Berger, Microsoft Azure and University of Washington

NCC: Natural Concurrency Control for Strictly Serializable Datastores by Avoiding the Timestamp-Inversion Pitfall ................................................................. 305
Haonan Lu, University at Buffalo; Shuai Mu, Stony Brook University; Siddhartha Sen, Microsoft Research; Wyatt Lloyd, Princeton University

Conveyor: One-Tool-Fits-All Continuous Software Deployment at Meta ...................................... 325
Boris Grubic, Meta; Yang Wang, Meta and the Ohio State University; Tyler Petrochko, Ran Yaniv, Brad Jones, David Callies, Matt Clarke-Lauer, and Dan Kelley, Meta; Soteris Demetriou, Meta and Imperial College London; Kenny Yu and Chunqiang Tang, Meta

Query Your Bits

Chardonnay: Fast and General Datacenter Transactions for On-Disk Databases ........................ 343
Tamer Eldeeb and Xincheng Xie, Columbia University; Philip A. Bernstein, Microsoft Research; Asaf Cidon and Junfeng Yang, Columbia University
Defcon: Preventing Overload with Graceful Feature Degradation

Cilantro: Performance-Aware Resource Allocation for General Objectives via Online Feedback
Romil Bhardwaj, UC Berkeley; Kirthasananda Kasandasamy, University of Wisconsin-Madison; Asim Biswal, Wenshuo Guo, Benjamin Hindman, Joseph Gonzalez, Michael Jordan, and Ion Stoica, UC Berkeley

Karma: Resource Allocation for Dynamic Demands
Midhul Vuppalapati, Giannis Fikioris, and Rachit Agarwal, Cornell University; Asaf Cidon, Columbia University; Anurag Khandelwal, Yale University; Éva Tardos, Cornell University

Wednesday, July 12

Train Your Bits I

AlpaServe: Statistical Multiplexing with Model Parallelism for Deep Learning Serving
Zhuohan Li and Lianmin Zheng, UC Berkeley; Yinmin Zhong, Peking University; Vincent Liu, University of Pennsylvania; Ying Sheng, Stanford University; Xin Jin, Peking University; Yanping Huang and Zhifeng Chen, Google; Hao Zhang, UC San Diego; Joseph E. Gonzalez and Ion Stoica, UC Berkeley

Cocktailer: Analyzing and Optimizing Dynamic Control Flow in Deep Learning
Chen Zhang, Tsinghua University; Lingxiao Ma and Jilong Xue, Microsoft Research; Yining Shi, Peking University & Microsoft Research; Ziming Miao and Fan Yang, Microsoft Research; Jidong Zhai, Tsinghua University; Zhi Yang, Peking University; Haojue Wang, Jidong Zhai, Muyan Hu, Zixuan Ma, Tuowei Wang, and Shuhong Huang, Tsinghua University; Xupeng Miao, Carnegie Mellon University; Shizhi Tang and Kezhao Huang, Tsinghua University; Zhihao Jia, Carnegie Mellon University

Welder: Scheduling Deep Learning Memory Access via Tile-graph.
Yining Shi, Peking University & Microsoft Research; Zhi Yang, Peking University; Jilong Xue, Lingxiao Ma, Yuqing Xia, Ziming Miao, Yuxiao Guo, Fan Yang, and Lidong Zhou, Microsoft Research

Effectively Scheduling Computational Graphs of Deep Neural Networks toward Their Domain-Specific Accelerators
Jie Zhao, Information Engineering University; Siyuan Feng, Shanghai Jiao Tong University; Xiaojing Dan, Fei Liu, Chengke Wang, Sheng Yuan, Wenyuan Lv, and Qikai Xie, Stream Computing Inc

Train Your Bits II

EinNet: Optimizing Tensor Programs with Derivation-Based Transformations
Liyan Zheng, Haojie Wang, Jidong Zhai, Muyan Hu, Zixuan Ma, Tuowei Wang, and Shuhong Huang, Tsinghua University; Xupeng Miao, Carnegie Mellon University; Shizhi Tang and Kezhao Huang, Tsinghua University; Zhiliang Jia, Carnegie Mellon University

Hydro: Surrogate-Based Hyperparameter Tuning Service in Datacenters
Qinghao Hu, Nanyang Technological University, S-Lab, NTU, and Shanghai AI Laboratory; Zhisheng Ye, Shanghai AI Laboratory and Peking University; Meng Zhang, Nanyang Technological University, S-Lab, NTU, and Shanghai AI Laboratory; Qiaoling Chen, Shanghai AI Laboratory and National University of Singapore; Peng Sun, Shanghai AI Laboratory and SenseTime Research; Yonggang Wen and Tianwei Zhang, Nanyang Technological University

MGG: Accelerating Graph Neural Networks with Fine-Grained Intra-Kernel Communication-Computation Pipelining on Multi-GPU Platforms
Yuke Wang, Boyuan Feng, and Zheng Wang, University of California Santa Barbara; Tong Geng, University of Rochester; Kevin Barker and Ang Li, Pacific Northwest National Laboratory; Yufei Ding, University of California Santa Barbara

Optimizing Dynamic Neural Networks with Brainstorm
Weihao Cui, Shanghai Jiao Tong University; Zhenhua Han, Microsoft Research Asia; Lingji Ouyang, University of Science and Technology of China; Yichuan Wang, Shanghai Jiao Tong University; Ningxin Zheng, Lingxiao Ma, Yuqing Yang, Fan Yang, Jilong Xue, Lili Qiu, and Lidong Zhou, Microsoft Research Asia; Quan Chen, Shanghai Jiao Tong University; Haisheng Tan, University of Science and Technology of China; Minyi Guo, Shanghai Jiao Tong University