

20th USENIX Symposium on Operating Systems Design and Implementation (OSDI '26)

July 13–15, 2026

Seattle, WA, USA

Monday, July 13

KV Cache and Long Context

Strata: Hierarchical Context Caching for Long Context Language Model Serving 1

Zhiqiang Xie, *Stanford University and NVIDIA*; Ziyi Xu, *Shanghai Jiao Tong University*; Mark Zhao, *University of Colorado Boulder*; Yuwei An, *Carnegie Mellon University*; Vikram Sharma Mailthody, *NVIDIA*; Scott Mahlke, *NVIDIA and University of Michigan*; Michael Garland, *NVIDIA*; Christos Kozyrakis, *NVIDIA and Stanford University*

ECHO: Efficient KV Cache Offloading with Lossless Prefetching for Serving Native Sparse Attention LLMs 17

Guangda Liu, Wenhao Chen, Chengwei Li, and Zhenyu Ning, *Shanghai Jiao Tong University*; Jing Lin and Yiwu Yao, *Huawei*; Quan Chen, Shixuan Sun, and Jieru Zhao, *Shanghai Jiao Tong University*; Minyi Guo, *Guizhou University and Shanghai Jiao Tong University*

No Buffer, No Bottleneck: Efficient Zero-Copy KV Cache Offloading for Long-Context LLMs 39

Shutian Luo and Haiying Shen, *University of Virginia*

Simple is Better: Multiplication May Be All You Need for LLM Request Scheduling 55

Dingyan Zhang, Jinbo Han, Kaixi Zhang, and Xingda Wei, *Shanghai Jiao Tong University*; Sijie Shen, Chenguang Fang, Wen Yuan Yu, and Jingren Zhou, *Alibaba Group*; Rong Chen, *Shanghai Jiao Tong University*

Prism: Cost-Efficient Multi-LLM Serving via GPU Memory Ballooning 75

Shan Yu, *University of California, Los Angeles*; Yifan Qiao, *University of California, Berkeley*; Mingyuan Ma, *Harvard University*; Yangmin Li, *Carnegie Mellon University*; Shuo Yang, *University of California, Berkeley*; Xinyuan Tong, *University of Edinburgh*; Yang Wang, *Intel*; Zhiqiang Xie, *Stanford University*; Yuwei An, *Carnegie Mellon University*; Shiyi Cao, *University of California, Berkeley*; Ke Bao, *LMSYS*; Deepak Vij, Xiaoning Ding, and Yichen Wang, *ByteDance*; Qingda Lu, *Alibaba Cloud*; Zhong Wang, *Tsinghua University*; Gao Gao, *Novita AI*; Harry Xu and Junyi Shu, *University of California, Los Angeles*; Jiarong Xing, *Rice University*; Ying Sheng, *University of California, Los Angeles*

Memory Tiering and CXL

Break on Through to the Other Side: Pooling Memory Elastically with RAMRYDER 95

Yanbo Zhou, *University of California, San Diego*; Erci Xu, *Shanghai Jiao Tong University*; Dongjoo Seo and Adam Manzanarez, *Samsung Semiconductor*; Steven Swanson, *University of California, San Diego*

MAC: Metadata Acceleration for Sustainable Performance in Big-Data Systems with CXL DRAM. 115

Dusol Lee, *Seoul National University*; Yan Sun, Houxiang Ji, Vinit Gupta, and Austin Antony Cruz, *University of Illinois Urbana–Champaign*; Inhyuk Choi, *Seoul National University*; Nam Sung Kim, *University of Illinois Urbana–Champaign*; Jihong Kim, *Seoul National University*

Finding NEMO: Nimble and Expressive Memory Observability 133

Shihang Li and Matthew Giordano, *University of Washington*; Tushar Garg, *Meta*; Rohan Kadekodi, *University of Washington*; Daniel S. Berger, *University of Washington and Microsoft Azure*; Baris Kasikci, Thomas Anderson, and Simon Peter, *University of Washington*

OBASE: Object-Based Address-Space Engineering to Improve Memory Tiering 151

Vinay Banakar, *University of Wisconsin–Madison and Google*; Suli Yang, *Google*; Kan Wu, *xAI*; Andrea C. Arpaci-Dusseau and Remzi H. Arpaci-Dusseau, *University of Wisconsin–Madison*; Kimberly Keeton, *Google*

MDK: Rethinking the data center memory reclamation problem 167

Shaurya Patel, *Google and University of British Columbia*; Suli Yang and Yawen Wang, *Google*; Kan Wu, *xAI*; Alexandra (Sasha) Fedorova, *University of British Columbia and MongoDB*; Margo Seltzer, *University of British Columbia*; Kimberly Keeton, *Google*

Sandboxing, Access Control, and Drivers

USEC: A User-Requirement-Driven Mandatory Access Control Framework for Operating Systems (Operational Systems)	191
Yu Jiang, <i>Tsinghua University</i> ; Wenhuan Liu, <i>Tsinghua University and UnionTech Software Technology Co., Ltd</i> ; Fuchen Ma, Yuheng Shen, and Yuanliang Chen, <i>Tsinghua University</i> ; Lei Zhang and He Li, <i>UnionTech Software Technology Co., Ltd.</i> ; Quan Zhang and Chijin Zhou, <i>East China Normal University</i>	
Mohabi: Disaggregating and Sandboxing the Firefox JavaScript Engine	207
Abhishek Sharma and Anand Balaji, <i>The University of Texas at Austin</i> ; Zachary Yedidia, <i>Stanford University</i> ; Anthony Du and Taehyun Noh, <i>The University of Texas at Austin</i> ; Iain Ireland, Jan de Mooij, and Matthew Gaudet, <i>Mozilla</i> ; Tal Garfinkel, <i>Google</i> ; Deian Stefan and Hovav Shacham, <i>University of California, San Diego</i> ; Shravan Narayan, <i>The University of Texas at Austin</i>	
Ichnaea: A Framework for Precise Tracking of Memory Objects	229
Samad Haque and Sibin Mohan, <i>The George Washington University</i> ; Aaron Paulos and Partha Pal, <i>RTX BBN Technologies</i>	
Extracting Database Access-Control Policies From Web Applications	249
Wen Zhang, Dev Bali, and Jamison Kerney, <i>University of California, Berkeley</i> ; Aurojit Panda, <i>NYU</i> ; Scott Shenker, <i>University of California, Berkeley, and ICSI</i>	
iLand: An Instruction-Level Dynamic Binary Instrumentation framework for iOS	271
Kaitao Xie, Yizhuo Wang, and Xiaolong Bai, <i>Alibaba Group</i>	

LLM Training at Scale

Tessera: A Holistic Pipeline Parallelism Framework for Trillion-Parameter Heterogeneous MoE Training (Operational Systems)	287
Weifang Hu, <i>Huazhong University of Science and Technology</i> ; Langshi Chen, Man Yuan, Youyang Yao, Xiulong Yuan, Li Tian, Yong Li, and Wei Lin, <i>Alibaba Cloud</i> ; Xuanhua Shi, <i>Huazhong University of Science and Technology</i> ; Zhengping Qian and Jingren Zhou, <i>Alibaba Cloud</i>	
Hetu v2: A General and Scalable Deep Learning System with Hierarchical and Heterogeneous Single Program Multiple Data Annotations	305
Haoyang Li, <i>Peking University</i> ; Fangcheng Fu, <i>Shanghai Jiao Tong University</i> ; Hao Ge, Sheng Lin, Xuanyu Wang, Jiawen Niu, and Yuming Zhou, <i>Peking University</i> ; Xupeng Miao, <i>Purdue University</i> ; Bin Cui, <i>Peking University and Peking University (Qingdao)</i>	
Syncopate: Efficient Multi-GPU AI Kernels via Automatic Chunk-Centric Compute-Communication Overlap ...	331
Xinwei Qiang, Yue Guan, and Zhengding Hu, <i>University of California, San Diego</i> ; Keren Zhou, <i>George Mason University and OpenAI</i> ; Yufei Ding, <i>University of California, San Diego, and Meta</i> ; Adnan Aziz, <i>Meta</i>	
Teaching The Old Dog New Tricks: Building Efficient Data Pipelines for Large-Scale LLM Pre-training (Operational Systems)	349
Luofan Chen and Chenhan Wang, <i>University of Science and Technology of China and ByteDance Seed</i> ; Weidong Zhang, Jinxin Chi, Hequan Zhang, Zhanbo Wang, Chenyuan Wang, Lishu Luo, Sijin Wu, Junqi Hu, Jun Wang, and Cheng Chen, <i>ByteDance Seed</i> ; Lixin Huang, Liyang Zhao, Yong Tian, and Jun Guo, <i>ByteDance</i> ; Youhui Bai, <i>University of Science and Technology of China</i> ; Wencong Xiao, <i>ByteDance Seed</i> ; Kang Chen, <i>Tsinghua University</i> ; Cheng Li, <i>University of Science and Technology of China and Institute of Artificial Intelligence, Hefei Comprehensive National Science Center</i>	
Cocoon: A System Architecture for Differentially Private Training with Correlated Noises	367
Donghwan Kim and Xin Gu, <i>The Pennsylvania State University</i> ; Jinho Baek, Timothy Lo, Younghoon Min, Kwangsik Shin, and Jongryool Kim, <i>SK hynix Inc.</i> ; Jongse Park, <i>Korea Advanced Institute of Science and Technology (KAIST)</i> ; Kiwan Maeng, <i>The Pennsylvania State University</i>	

Testing, Debugging, and Root Causing

ValScope: Value-Semantics-Aware Metamorphic Testing for Detecting Logical Bugs in DBMSs	387
Li Lin, Liehang Chen, and Rongxin Wu, <i>Xiamen University</i>	
The Abstention Protocol: RCA for Clos Fabrics (Operational Systems)	405
Madhava Gaikwad, <i>Independent</i> ; and Deepak Pandey, <i>Microsoft</i>	

When Sampling Lies: Trustworthy Performance Profiling for Flat Workloads with Blink (Operational Systems)	423
Rishikesh Devsot, <i>YScope</i> ; ChenXing Yang and Yi Fan Yu, <i>YScope and University of Toronto</i> ; Prabhdeep Singh Soni, Afshin Arefi, Bryan Chan, and Reza Azimi, <i>Huawei Technologies Canada</i> ; Ding Yuan, <i>YScope and University of Toronto</i>	
Breaking the Reward Barrier: Accelerating Tree-of-Thought Reasoning via Speculative Exploration	437
Shuzhang Zhong, Haochen Huang, and Shengxuan Qiu, <i>Peking University</i> ; Pengfei Zuo, <i>ByteDance Seed</i> ; Runsheng Wang and Meng Li, <i>Peking University</i>	
Controlling Opaque-Component Effects with Semisolates and Try	453
Evangelos Lamprou, <i>Brown University</i> ; Tianyu (Ezri) Zhu, <i>Stevens Institute of Technology</i> ; Di Jin and Grigoris Ntousakis, <i>Brown University</i> ; Georgios Liargkovas, <i>Columbia University</i> ; Calvin Eng, <i>Brown University</i> ; Konstantinos Kallas, <i>University of California, Los Angeles</i> ; Michael Greenberg, <i>Stevens Institute of Technology</i> ; Nikos Vasilakis, <i>Brown University</i>	
Kernel Scheduling and Tail Latency	
SBB: Eliminating Centralized Bottlenecks in Userspace Network Runtime	473
Kang Hu, <i>Peking University and Zhongguancun Laboratory</i> ; Shuqi Dong, <i>Peking University</i> ; Chuandong Li, <i>Peking University and Zhongguancun Laboratory</i> ; Ran Yi, <i>Peking University</i> ; Zonghao Zhang, <i>Peking University and Zhongguancun Laboratory</i> ; Yiming Yao, <i>Peking University</i> ; Bo An, <i>Zhongguancun Laboratory</i> ; Jie Zhang, Xiaolin Wang, and Yingwei Luo, <i>Peking University and Zhongguancun Laboratory</i> ; Zhenlin Wang, <i>Michigan Technological University</i> ; Diyu Zhou, <i>Peking University</i>	
Rakaia: Scalable In-Kernel Scheduling for TCP-Based RPCs	491
Rui Yang, Konstantinos Prasopoulos, and Edouard Bugnion, <i>EPFL</i>	
kSTEP: Characterization and Deterministic Testing of Linux CPU Scheduler Bugs	509
Tingjia Cao, Shawn (Wanxiang) Zhong, and Caeden Whitaker, <i>University of Wisconsin–Madison</i> ; Ke Han, <i>Purdue University</i> ; Andrea C. Arpaci-Dusseau and Remzi H. Arpaci-Dusseau, <i>University of Wisconsin–Madison</i>	
What Are You (M)Waiting For: The Hidden Cost of Idle in the Hyperscale Cloud (Operational Systems)	529
Yun Wang, <i>Shanghai Jiao Tong University</i> ; Xingguo Jia, <i>Alibaba Cloud</i> ; Ben Luo and Kenan Liu, <i>Alibaba Group</i> ; Shengdong Dai, <i>Alibaba Cloud</i> ; Jingdong Han and Weihao Chen, <i>Alibaba Group</i> ; Yicheng Gu and Xingzi Yu, <i>Shanghai Jiao Tong University</i> ; Yibin Shen and Jiesheng Wu, <i>Alibaba Cloud</i> ; Zhengwei Qi and Haibing Guan, <i>Shanghai Jiao Tong University</i>	
Xkernel: Principled Performance Tunability of Operating System Kernels	547
Zhongjie Chen, <i>Tsinghua University and Microsoft Research</i> ; Wentao Zhang, <i>University of Illinois Urbana–Champaign</i> ; Yulong Tang and Ran Shu, <i>Microsoft Research</i> ; Fengyuan Ren, <i>Tsinghua University</i> ; Tianyin Xu, <i>University of Illinois Urbana–Champaign</i> ; Jing Liu, <i>Microsoft Research</i>	
Agentic AI and LLM Operations	
Murakkab: Resource-Efficient Agentic Workflow Orchestration in Cloud Platforms	567
Gohar Irfan Chaudhry, <i>MIT CSAIL</i> ; Esha Choukse, Haoran Qiu, Íñigo Goiri, and Rodrigo Fonseca, <i>Microsoft Azure Research</i> ; Adam Belay, <i>MIT CSAIL</i> ; Ricardo Bianchini, <i>Microsoft Azure</i>	
ECO: An AI-Driven Code Efficiency Optimizer for Warehouse Scale Computers (Operational Systems)	589
Hannah Lin and Martin Maas, <i>Google DeepMind</i> ; Maximilian Roquemore, <i>Google</i> ; Arman Hasanzadeh, <i>Google DeepMind</i> ; Fred Lewis, Yusuf Simonson, Ameya Shringi, and Hongwen Dai, <i>Google</i> ; Patrick Musau, <i>Google DeepMind</i> ; Tzu-Wei Yang, <i>Google</i> ; Amir Yazdanbakhsh and Deniz Altinbüken, <i>Google DeepMind</i> ; Florin Papa, Maggie Nolan Edmonds, Aditya Patil, Don Schwarz, Satish Chandra, and Chris Kennelly, <i>Google</i> ; Milad Hashemi, <i>Google DeepMind</i> ; Parthasarathy Ranganathan, <i>Google</i>	
StriaTrace: Efficient Tracing and Diagnosis for Online LLM Inference (Operational Systems)	627
Haonan Wu, <i>Shanghai Jiao Tong University and Alibaba Group</i> ; Yanqing Chen, Kun Qian, Xue Li, and Jingbo Xu, <i>Alibaba Group</i> ; Erci Xu, <i>Shanghai Jiao Tong University</i> ; Ennan Zhai and Wenyuan Yu, <i>Alibaba Group</i> ; Guangtao Xue, <i>Shanghai Jiao Tong University and Shanghai Key Laboratory of Trusted Data Circulation and Governance and Web3</i> ; Jingren Zhou, <i>Alibaba Group</i>	
Diagnosing Performance Issues in Application-Defined Resources	647
Yigong Hu and You-Liang Huang, <i>Boston University</i> ; Haodong Zheng, <i>University of Washington and EPFL</i> ; Yicheng Liu, <i>University of Washington and University of California, Los Angeles</i> ; Dedong Xie and Baris Kasikci, <i>University of Washington</i>	

Program Analysis

- hS: Speculative Script Reordering at Subprocess Granularity** 665
Georgios Liargkovas and Di Jin, *Brown University*; Tianyu (Ezri) Zhu and Dan Liu, *Stevens Institute of Technology*;
A. Bolun Thompson, *University of California, Los Angeles*; Anirudh Narsipur, Seong-Heon Jung, and Siddhartha Prasad,
Brown University; Diomidis Spinellis, *Athens University of Economics and Business and TU Delft*; Michael Greenberg,
Stevens Institute of Technology; Konstantinos Kallas, *University of California, Los Angeles*;
Nikos Vasilakis, *Brown University*
- Incr: Faster Re-execution via Bolt-on Incrementalization** 683
Yizheng Xie, Evangelos Lamprou, Jerry Xia, and Nikos Vasilakis, *Brown University*
- A Compilation-based Under-Constrained Execution Engine** 701
Mingjun Yin, Zhaorui Li, Ju Chen, Haochen Zeng, and Chengyu Song, *University of California, Riverside*
- Aletheia: Automated Detection of Data Integrity Violations in Microservices** 721
Mafalda Sofia Ferreira, João Ferreira Loff, João Garcia, and Rodrigo Rodrigues, *INESC-ID, Instituto Superior Técnico, Universidade de Lisboa*

Synchronization

- ARCTIC: a practical lock-free adaptive radix tree** 739
Newton Ni, Nicolas Garza, Jenny Stinehour, and Michael Goppert, *The University of Texas at Austin*;
Michal Friedman, *ETH Zürich*; Emmett Witchel, *The University of Texas at Austin*
- Efficient and Scalable Synchronization via Generalized Cache Coherence** 755
Yanpeng Yu, Seung-seob Lee, Lin Zhong, and Anurag Khandelwal, *Yale University*
- Shaving the Peaks: Taming Tail Latency for Managed Workloads via Disaggregated Garbage Collection** 773
Hongtao Lyu, Yuhan Li, and Mingyu Wu, *Shanghai Jiao Tong University*
- DeLFS: A Decentralized Log-Structured File System for Manycores** 793
Taehwan Ahn, Chanhyeong Yu, Sangjin Lee, and Yongseok Son, *Chung-Ang University*

Tuesday, July 14

RL Training at Scale

- Weave: Efficient Co-Scheduling for Disaggregated RL Post-Training** 809
Tianyuan Wu and Lunxi Cao, *Hong Kong University of Science and Technology*; Yining Wei, *University of Illinois Urbana-Champaign*; Wei Gao, Yuheng Zhao, and Dakai An, *Hong Kong University of Science and Technology*;
Shaopan Xiong, Zhiqiang Lv, Ju Huang, Siran Yang, Yinghao Yu, Jiamang Wang, and Lin Qu, *Alibaba Group*;
Wei Wang, *Hong Kong University of Science and Technology*
- RLinf: Flexible and Efficient Large-Scale Reinforcement Learning via Macro-to-Micro Flow Transformation** ... 829
Chao Yu, *Tsinghua University*; Yuanqing Wang, *Infinigence AI and Peking University*; Zhen Guo, Hao Lin, and Si Xu, *Infinigence AI*; Hongzhi Zang, *Tsinghua University*; Quanlu Zhang, *Infinigence AI*; Yongji Wu, *University of California, Berkeley*; Chunyang Zhu and Junhao Hu, *Infinigence AI*; Zixiao Huang, *Tsinghua University and Infinigence AI*; Mingjie Wei, *Zhongguancun Academy*; Yuqing Xie, *Tsinghua University*; Ke Yang, *Zhongguancun Academy*; Bo Dai, *Beihang University and Infinigence AI*; Zhexuan Xu and Jiakun Du, *Tsinghua University*; Xiangyuan Wang, *Peking University and Infinigence AI*; Xu Fu and Letong Shi, *Infinigence AI*; Zhihao Liu, *Zhongguancun Academy*; Kang Chen, *Peking University and Zhongguancun Academy*; Weilin Liu, *Infinigence AI*; Gang Liu, *Tsinghua University*; Boxun Li, *Infinigence AI*; Jianlei Yang, *Beihang University*; Zhi Yang, *Peking University*; Guohao Dai, *Shanghai Jiao Tong University and Infinigence AI*; Yu Wang, *Tsinghua University*
- DynaRL: Flexible and Dynamic Scheduling of Large-Scale Reinforcement Learning Training** 847
Yuanqing Wang, *Peking University and Infinigence AI*; Hao Lin, Junhao Hu, Chunyang Zhu, Quanlu Zhang, and Zhen Guo, *Infinigence AI*; Yuchen Zhang, *Institute of Computing Technology, Chinese Academy of Sciences and Infinigence AI*; Xu Fu and Si Xu, *Infinigence AI*; Bo Dai, *Beihang University and Infinigence AI*; Zixiao Huang, *Tsinghua University and Infinigence AI*; Chao Yu, *Tsinghua University*; Boxun Li, *Infinigence AI*; Guohao Dai, *Shanghai Jiao Tong University and Infinigence AI*; Zhi Yang, *Peking University*; Yu Wang, *Tsinghua University*

ROLLART: Disaggregated Multi-Task Agentic RL Training at Scale. 863
 Wei Gao, Yuheng Zhao, and Tianyuan Wu, *Hong Kong University of Science and Technology*; Shaopan Xiong and Weixun Wang, *Alibaba Group*; Dakai An and Lunxi Cao, *Hong Kong University of Science and Technology*; Dilxat Muhtar, Zichen Liu, Haizhou Zhao, Ju Huang, and Siran Yang, *Alibaba Group*; Yongbin Li, *Tongyi Lab, Alibaba*; Wenbo Su, Jiamang Wang, Lin Qu, and Bo Zheng, *Alibaba Group*; Wei Wang, *Hong Kong University of Science and Technology*

Seer: Online Context Learning for Fast Synchronous LLM Reinforcement Learning. 883
 Ruoyu Qin, *Moonshot AI and Tsinghua University*; Weiran He, Weixiao Huang, Yangkun Zhang, Yikai Zhao, Bo Pang, and Xinran Xu, *Moonshot AI*; Yingdi Shan, Yongwei Wu, and Mingxing Zhang, *Tsinghua University*

Disaggregated Memory Systems

Harvesting Sub-Microsecond CXL Memory Stalls with LiteSwitch. 903
 Nanqinqin Li, *Princeton University*; Yuhong Zhong and Asaf Cidon, *Columbia University*; Michael J. Freedman, *Princeton University*

Duhu: Shared Disaggregated Memory for Distributed Data Processing Frameworks. 921
 Qiutong Men and Tao Wang, *New York University*; Jongryool Kim and Hane (Stella) Yie, *SK hynix*; Emmanuel Amaro, *Microsoft*; Marcos K. Aguilera, *NVIDIA*; Aurojit Panda, *New York University*

Blowfish: Elastic Virtual Machine Memory for Disaggregated Memory 939
 Yulong Zhang, *SKLP, Institute of Computing Technology, CAS and University of Chinese Academy of Sciences*; Yilong Luo and Diyu Zhou, *SCS, Peking University, China*; Quan Chen, *Shanghai Jiao Tong University*; Quanxi Li, *SKLP, Institute of Computing Technology, CAS and University of Chinese Academy of Sciences*; Mosong Zhou, Lei Zhu, Senbo Fu, and Qian Peng, *Huawei Cloud*; Huimin Cui and Xiaobing Feng, *SKLP, Institute of Computing Technology, CAS and University of Chinese Academy of Sciences*; Tao Xie, *SCS, Peking University, China*; Chenxi Wang, *SKLP, Institute of Computing Technology, CAS and University of Chinese Academy of Sciences*

Espresso: Constructing Cost-Efficient CXL JBOF via Inter-SSD Computing Resource Sharing 957
 Shushu Yi, Yuda An, Li Peng, and Xiurui Pan, *Peking University*; Qiao Li, *Mohamed bin Zayed University of Artificial Intelligence*; Jieming Yin, *Nanjing University of Posts and Telecommunications*; Guangyan Zhang, *Tsinghua University*; Wenfei Wu, *Peking University*; Chenxi Wang, *University of Chinese Academy of Sciences*; Diyu Zhou, *Peking University*; Zhenlin Wang, *Michigan Tech*; Xiaolin Wang and Yingwei Luo, *Peking University*; Ke Zhou, *Huazhong University of Science and Technology (HUST)*; Jie Zhang, *Peking University*

FORGE: Mitigating Synchronization Amplification for Memory-Disaggregated Caching Systems. 977
 Zhijun Yang, Yu Hua, Ming Zhang, Menglei Chen, and Yixiao Wang, *Huazhong University of Science and Technology*

Confidential Computing

Accelerating Confidential Databases with Crypto-free Mappings 997
 Wenxuan Huang, Zhanbo Wang, and Mingyu Li, *Key Laboratory of System Software, Chinese Academy of Sciences, and Institute of Software, Chinese Academy of Sciences, and University of Chinese Academy of Sciences*

JANUS: Cross-World, Cooperative Nested Virtualization for Secure Containers 1015
 Jiangshan Lai, *Ant Group*; Hang Huang, *Alibaba Cloud and Huazhong University of Science and Technology*; Quan Xu and Zhen Ren, *Alibaba Cloud*; Wenlong Hou, *Ant Group*; Wei Guo, *Alibaba Cloud*; Jia Rao and Hui Lu, *The University of Texas at Arlington*; Weidong Han, Jiesheng Wu, Jiang Liu, Naixuan Guan, and Yibin Shen, *Alibaba Cloud*; Feng Yu and Xu Wang, *Ant Group*; Shiqiang Zhang, *Alibaba Cloud*; Zhiheng Tao, *Ant Group*; Yisheng Xie, *Alibaba Cloud*; Song Wu and Hai Jin, *Huazhong University of Science and Technology*

Osprey: Transparent and Efficient Virtual Memory for Secure Computation 1033
 Yicheng Liu, *University of California, Los Angeles*; Alice Yeh, *University of California, Berkeley*; Harry Xu, *University of California, Los Angeles*; Raluca Ada Popa, *University of California, Berkeley*; Sam Kumar, *University of California, Los Angeles*

Nested SEV: Secure and Generic SEV Support for Nested Virtualization. 1051
 Kazuki Takiguchi and Kenichi Kourai, *Kyushu Institute of Technology*

μUSB: Practical and Safe USB Driver Reuse for Arm TrustZone 1067
Xunkai Zhang, Sijin Li, and Pei Meng, *University of Electronic Science and Technology of China*; Meng Wang, *CISPA Helmholtz Center for Information Security*; Yongzhao Zhang, Ting Chen, Xiaosong Zhang, and Liwei Guo, *University of Electronic Science and Technology of China*

Expert Mixture

Achieving Cloud-Grade SLOs for Local Mixture-of-Experts Inference through CPU–GPU Hybrid Design 1089
Wenxin Wang, *Tsinghua University*; Yule Hou and Yu Ji, *Xingyun Integrated Circuits Co., Ltd.*; Peng Qu and Youhui Zhang, *Tsinghua University and Beijing National Research Center for Information Science and Technology*

UEP: Portable Expert-Parallel Communication 1107
Ziming Mao, *University of California, Berkeley*; Yihan Zhang, *University of California, Davis*; Chihan Cui, *University of Wisconsin–Madison*; Zhen Huang, *AMD*; Kaichao You, *Independent Researcher*; Zhongjie Chen, *Tsinghua University*; Zhiying Xu, *Amazon Web Services*; Zhenyu Gu, *AMD*; Scott Shenker, *University of California, Berkeley, and ICSI*; Costin Raiciu, *Broadcom and Politehnica of Bucharest*; Yang Zhou, *University of California, Davis*; Ion Stoica, *University of California, Berkeley*

BatchGen: An Architecture for Scalable and Efficient Batch Inference 1125
Tairan Xu, Leyang Xue, and Zhan Lu, *University of Edinburgh*; Jinfu Deng and Hongyang Xiao, *Tencent*; Yinsicheng Jiang, Congjie He, Matej Sandor, Le Xu, and Luo Mai, *University of Edinburgh*

UCCL-Tran: An Extensible Software Transport Layer for GPU Networking 1143
Yang Zhou, *University of California, Berkeley, and University of California, Davis*; Zhongjie Chen, *Tsinghua University*; Ziming Mao, *University of California, Berkeley*; ChonLam Lao, *Harvard University*; Shuo Yang, *University of California, Berkeley*; Pravein Govindan Kannan, *IBM Research*; Xizhi Zhang, *Tsinghua University*; Jiaqi Gao, *Independent Researcher*; Yilong Zhao and Yongji Wu, *University of California, Berkeley*; Kaichao You, *Independent Researcher*; Fengyuan Ren, *Tsinghua University*; Zhiying Xu, *Amazon Web Services*; Costin Raiciu, *Broadcom and University Politehnica of Bucharest*; Ion Stoica, *University of California, Berkeley*

Power, Energy, and Sustainability

Hardware Lifecycle-Aware Power Planning in Commercial Hyperscale Datacenters (Operational Systems) 1167
Ruihao Li, *Meta and The University of Texas at Austin*; Leonardo Piga, Wei Su, and Carlos Torres, *Meta*; Jovan Stojkovic, *Meta and The University of Texas at Austin*; Neeraja J. Yadwadkar and Lizy K. John, *The University of Texas at Austin*; Abhishek Dhanotia, *Meta*

Kareus: Joint Reduction of Dynamic and Static Energy in Large Model Training 1185
Ruofan Wu, Jae-Won Chung, and Mosharaf Chowdhury, *University of Michigan*

SPADE: Signal-Aware DAG Scheduling and Dynamic Provisioning for Data Processing Clusters 1205
Adam Lechowicz, *University of Massachusetts Amherst*; Rohan Shenoy, *University of California, Berkeley*; Noman Bashir, *Massachusetts Institute of Technology*; Mohammad Hajiesmaili, *University of Massachusetts Amherst*; Adam Wierman, *California Institute of Technology*; Christina Delimitrou, *Massachusetts Institute of Technology*

Quota Marketplace: Dynamic Pricing for Efficient Allocation of ML Training Resources. 1223
Balasubramanian Sivan, Renato Paes Leme, Mihai Tiuca, and Ian McFarlane, *Google*; Vasilis Gkatzelis, *Google and Drexel University*; Nehal Mehta, Soheil Hassas Yeganeh, Vahab Mirrokni, and Amin Vahdat, *Google*

Consensus and Byzantine Fault Tolerance

BODEGA: Localized Linearizable Reads at Anywhere Anytime via Roster Leases 1243
Guanzhou Hu, *Amazon Web Services*; Andrea C. Arpaci-Dusseau and Remzi H. Arpaci-Dusseau, *University of Wisconsin–Madison*

Equal Opportunity: A Correctness Condition for Ordered Consensus 1283
Yunhao Zhang, *Cornell University*; Haobin Ni, *University of Washington*; Soumya Basu, *OpenReserve Holdings*; Shir Cohen, *Cornell University*; Maofan Yin, *University of California, Santa Barbara*; Lorenzo Alvisi and Robbert van Renesse, *Cornell University*; Qi Chen and Lidong Zhou, *Microsoft Research*

Jetpack: Consensus Made Generally Fast	1299
Ze Tang, Zihao Zhang, and Weihai Shen, <i>Stony Brook University</i> ; Jicheng Shi, <i>DatenLord</i> ; Shuai Mu, <i>Stony Brook University</i>	
Ambulance: saving BFT through racing	1325
Neil Giridharan and Shubham Mishra, <i>University of California, Berkeley</i> ; Lorenzo Alvisi, <i>Cornell University</i> ; Natacha Crooks, <i>University of California, Berkeley</i> ; Benjamin Marsh, <i>Sei Labs</i> ; Hein Meling, <i>University of Stavanger</i> ; Kartik Nayak, <i>Duke University</i> ; Grzegorz Prusak, <i>Sei Labs</i>	
Training Reliability and Silent Errors	
SDCs in the Wild: Characterizing and Diagnosing SDC-defective GPUs in Production LLM Training (Operational Systems)	1349
Wenxin Zheng, <i>Shanghai Jiao Tong University and ByteDance Seed</i> ; Wenxiao Wang, Yun Zhang, and Mingcong Han, <i>ByteDance Seed</i> ; Bin Xu, Jinyu Gu, Xingda Wei, and Haibo Chen, <i>Shanghai Jiao Tong University</i> ; Zuquan Song, Gaohong Liu, Yucheng Nie, Zhe Nan, Zhuolin Zheng, Huan Yu, Shuguang Wang, Ziming Zhou, Hang Zhu, Wencong Xiao, and Xin Liu, <i>ByteDance Seed</i>	
Safeguarding LLM Training at Scale: Online SDC Detection and Insights from 35 Million GPU Hours	1369
Kinman Lei, <i>Tsinghua University</i> ; Liyan Zheng, Xiang Li, Hongmin Chen, Yun Zhang, Gaohong Liu, Zuquan Song, and Zixuan Ma, <i>ByteDance</i> ; Zhiyu Xue, <i>Tsinghua University</i> ; Minghui Yu, Shuguang Wang, Wencong Xiao, and Haibin Lin, <i>ByteDance</i> ; Yuyang Jin and Jidong Zhai, <i>Tsinghua University</i> ; Bo Liu and Xin Liu, <i>ByteDance</i>	
OPGUARD: Bitwise Alignment for Precise and General Debugging of Production LLM Training	1385
Ziming Zhou and Yinjie Zhao, <i>University of Michigan</i> ; Hang Zhu, Wenxiao Wang, Zhihao Bai, Yun Zhang, Shuguang Wang, and Haibin Lin, <i>ByteDance Seed</i> ; Peng Huang, <i>University of Michigan</i>	
RobustRL: Role-based Fault Tolerance System for RL Post-Training	1407
Zhenqian Chen and Baoquan Zhong, <i>Zhejiang University</i> ; Xiang Li, <i>unaffiliated</i> ; Qing Dai, Xinkui Zhao, and Miao Ye, <i>Zhejiang University</i> ; Ren Cheng, <i>unaffiliated</i> ; Lufei Zhang, <i>State Key Laboratory of Mathematical Engineering and Advanced Computing, China</i> ; Jianwei Yin, <i>Zhejiang University</i>	
Filesystems and I/O Acceleration	
Oxbow: A Coordinated Architecture for Multi-component File Systems	1425
Jongyul Kim, <i>University of Illinois Urbana–Champaign</i> ; Jaehwan Lee and Inhoe Koo, <i>KAIST</i> ; Peizhe Liu and Jiyuan Zhang, <i>University of Illinois Urbana–Champaign</i> ; Junho Ahn, <i>KAIST</i> ; Tianyin Xu, <i>University of Illinois Urbana–Champaign</i> ; Youngjin Kwon, <i>KAIST</i>	
Scaling the IO wall with Declarative IO	1443
Sanjith Athlur, <i>Carnegie Mellon University</i> ; Sara McAllister, <i>Carnegie Mellon University, University of Wisconsin, Madison, and Google</i> ; Theo Gregersen, Timothy Kim, Yiwei Chen, Sarvesh Tandon, and Lucy Wang, <i>Carnegie Mellon University</i> ; Daniel S. Berger, <i>Carnegie Mellon University, Microsoft Azure, and University of Washington</i> ; Saurabh Kadekodi and Arif Merchant, <i>Google</i> ; Benjamin Berg, <i>University of North Carolina at Chapel Hill</i> ; Nathan Beckmann, Rashmi Vinayak, George Amvrosiadis, and Gregory R. Ganger, <i>Carnegie Mellon University</i>	
Umap: Revisiting Memory-mapped I/O on Distributed File Systems for Efficient Matrix Access (Operational Systems)	1461
Yongchao He, <i>unaffiliated</i> ; Guangyan Zhang, <i>Tsinghua University</i> ; Zane Cao, <i>ScitiX AI</i> ; Wenfei Wu, <i>Peking University</i>	
CoPilotIO: CPU as a Co-pilot for GPU I/O to Free GPU Compute	1479
Guanyi Chen and Qi Chen, <i>The Hong Kong University of Science and Technology (Guangzhou)</i> ; Shu Yin, <i>ShanghaiTech University</i> ; Jian Zhang, <i>The Hong Kong University of Science and Technology (Guangzhou)</i>	
Close to the Network	
RoCE BALBOA: Service-Enhanced RDMA Offload Engine for Data Center SmartNICs	1495
Maximilian Jakob Heer, Benjamin Ramhorst, Yu Zhu, Luhao Liu, Zhiyi Hu, Jonas Dann, and Gustavo Alonso, <i>ETH Zurich</i>	
DPA-Store: An Ordered Network Data Path Key-Value Store	1513
Frederic Schimmelpfennig, <i>Johannes Gutenberg-Universität Mainz</i> ; Jan Sass, <i>Saarland University</i> ; Reza Salkhordeh, <i>Johannes Gutenberg-Universität Mainz</i> ; Martin Kröning and Stefan Lankes, <i>RWTH Aachen University</i> ; André Brinkmann, <i>Saarland University</i>	

FARLock: Asymmetric RDMA Locking Made Fair.	1531
Yuehao Hu, Jiatang Zhou, Tianzheng Wang, and Keval Vora, <i>Simon Fraser University</i>	
When DDIO Meets Page Coloring: Revisiting DDIO Performance with Sepia	1547
Changwoo Song, Sanghyun Kim, and Jinhyeok Oh, <i>Sungkyunkwan University</i> ; Qizhe Cai, <i>University of Virginia</i> ; Joonsung Kim and Jaehyun Hwang, <i>Sungkyunkwan University</i>	
Graphs, ANN, and Vector Search	
Disentangling Graph Dependencies for Efficient Billion-Scale GPU Vector Search.....	1563
Haoru Zhao, Jingkai He, Jingyao Zeng, Mingkai Dong, and Dong Du, <i>Shanghai Jiao Tong University</i>	
Efficient GPU-Centric Evolving Graph Processing at Scale	1585
Yunmo Zhang and Jiacheng Huang, <i>City University of Hong Kong</i> ; Xizhe Yin, <i>Independent Researcher</i> ; Junqiao Qiu, <i>City University of Hong Kong</i> ; Hong Xu, <i>The Chinese University of Hong Kong</i> ; Chun Jason Xue, <i>Mohamed bin Zayed University of Artificial Intelligence</i>	
Pluto: High-Performance, Memory-Efficient Distributed Graph Analytics Through Advanced Mirroring	1605
Ying-Wei Wu, <i>The University of Texas at Austin</i> ; Christopher J. Rossbach, <i>The University of Texas at Austin and Microsoft</i> ; Mattan Erez, <i>The University of Texas at Austin</i>	
The Clustering Strikes Back: Building Cost-Effective and High-Performance ANNS at Scale with Helmsman (Operational Systems).....	1623
Yuchen Huang and Baiteng Ma, <i>East China Normal University and Xiaohongshu Inc</i> ; Yiping Sun, Yang Shi, Xiao Chen, Xiaocheng Zhong, Zhiyong Wang, and Yao Hu, <i>Xiaohongshu Inc</i> ; Erci Xu, <i>Shanghai Jiaotong University</i> ; Chuliang Weng, <i>East China Normal University</i>	
Durable and Trustworthy Storage	
WiseCode: Breaking the Scalability Barriers of Wide-Stripe Vector Codes	1641
Sijie Cai, Guangyan Zhang, and Xiao Niu, <i>Tsinghua University</i>	
The LogDrive: Composable Durability for Cloud-Based Shared Logs.....	1663
Gardner Vickers, Lucas Bradstreet, Mahesh Balakrishnan, Prince Mahajan, David Mao, Xavier Léauté, Ismael Juma, Nikhil Bhatia, Jack Vanlightly, Prateek Jindal, Sumit Arrawatia, Andrew Grant, Dhruvil Shah, Dimitar Dimitrov, Gaurav Badoni, Shimiao Zhang, and Yang Yu, <i>Confluent</i>	
Timelock Drive: Isolated Time-Based Defense for Storage Systems	1683
Jonah Rosenblum, Juechu Dong, Peter Chen, and Satish Narayanasamy, <i>University of Michigan</i>	
High Fidelity Models for Large Scale Stateful Services (Operational Systems).....	1699
Nouraldin Jaber, Dongyun Jin, Bernhard Kragl, Enrico Magnago, Gustavo Petri, Thorsten Tarrach, and Serdar Tasiran, <i>Amazon Web Services</i>	
Virtualization and Live Migration	
M3U: Scalable Kernel Memory Management for Efficient Post-copy Live Migration of High-end Virtual Machines	1715
Yizhe Xu, <i>Shanghai Jiao Tong University</i> ; Yuan Tao, Zhibin Zhang, Kang Yan, and Chao Zhang, <i>Alibaba Cloud Computing</i> ; Shuo Shi, Zongpu Zhang, and Xu Huan, <i>Shanghai Jiao Tong University</i> ; Yibin Shen, Xudong Zheng, and Jiesheng Wu, <i>Alibaba Cloud Computing</i> ; Jian Li and Haibing Guan, <i>Shanghai Jiao Tong University</i>	
Compaction-Free Memory Defragmentation for Virtualization via Infinite Guest Physical Address Space	1731
Peixin Zeng, Hao Huang, Yanqi Pan, Wen Xia, Darong Yang, Jiahao Chen, and Nan Zhang, <i>Harbin Institute of Technology, Shenzhen</i>	
Inside Out: A Paradigm Shift In VM Introspection	1749
Dufy Tegua, <i>University Grenoble Alpes, CNRS, Inria, Grenoble INP, LIG, and Orange Research</i> ; Louis Duval, <i>University Grenoble Alpes, CNRS, Inria, Grenoble INP, LIG</i> ; Teo Pisenti, <i>IRIT, Université de Toulouse, CNRS, Toulouse INP, UT3 Toulouse</i> ; Kahina Lazri, <i>Orange Research</i> ; Daniel Hagimont, <i>IRIT, Université de Toulouse, CNRS, Toulouse INP, UT3 Toulouse</i> ; Thomas Pasquier, <i>University of British Columbia</i> ; Renaud Lachaize and Alain Tchana, <i>University Grenoble Alpes, CNRS, Inria, Grenoble INP, LIG</i>	

vBOIDS: Taming Chaos via Coarse-grained Scheduling Abstraction for Containers 1769
Kaesi Manakkal, *The University of Texas at Arlington*; Nathan Daughety, *Air Force Research Laboratory (AFRL)*;
Yu Sun, *Binghamton University*; Marcus Pendleton, *Air Force Research Laboratory (AFRL)*; Hui Lu, *The University of Texas at Arlington*

Wednesday, July 15

Resource-Efficient LLM Serving

Efficient LLM Serving on Commodity GPU Clusters with Data-Reduced Cross-Instance Orchestration 1787
Jiangsu Du, Hongbin Zhang, Taosheng Wei, Zhenyi Zheng, Jiazi Jiang, Kaiyi Wu, Zhiguang Chen, and Yutong Lu,
School of Computer Science and Engineering, Sun Yat-Sen University

Revisiting Pipeline Parallelism for LLM Serving 1803
Soonjae Hwang and Jeongseob Ahn, *Korea University*

OpenTela: Unifying Decentralized Computing Resources for Heterogeneous LLM Serving (Operational Systems) 1821
Xiaoze Yao, *ETH Zurich*; Youhe Jiang, *University of Cambridge*; Ilia Badanin, *EPFL*; Qinghao Hu, *MIT*;
Robert Matthew Smith, *ETH AI Center*; Binhang Yuan, *The Hong Kong University of Science and Technology*;
Imanol Schlag, *ETH AI Center*; Eiko Yoneki, *University of Cambridge*; Ana Klimovic, *ETH Zurich*

KAIROX: Adaptive GPU-CPU Hybrid LLM Inference via Online Neuron Balancing 1839
Yapeng Jiang, *Sun Yat-sen University, Peng Cheng Laboratory, and Zhuhai Key Laboratory of Trusted Large Language Models*; Minghao Gan, *Sun Yat-sen University*; Zicong Hong, *École Polytechnique Fédérale de Lausanne*; Wuhui Chen, *Sun Yat-sen University and Peng Cheng Laboratory*; Junyuan Liang, *Sun Yat-sen University*; Yue Yu, *Peng Cheng Laboratory*; Meng Guo, *Qilu University of Technology*; Zibin Zheng, *Sun Yat-sen University*

ADAngel: Accelerating Arbitrary-Precision Quantized LLMs with Adaptive Computing Mapping 1857
Yao Liu, Wenjie Wang, Yifei Feng, Bo Peng, Jianguo Yao, and Haibing Guan, *Shanghai Jiao Tong University*

GPU Compilers and Kernels

Optimal Software Pipelining and Warp Specialization for Tensor Core GPUs 1875
Rupanshu Soi, Rohan Yadav, Fredrik Kjolstad, and Alex Aiken, *Stanford University*; Maryam Mehri Dehnavi, Michael Garland, and Michael Bauer, *NVIDIA*

TileLoom: Automatic Dataflow Planning for Tile-Based Languages on Spatial Dataflow Accelerators 1891
Wei Li, Zhenyu Bai, Heru Wang, Pranav Dangi, and Zhiqiang Zhang, *National University of Singapore*; Cheng Tan, *Arizona State University and Google*; Huiying Lan, *Lumai Ltd.*; Weng-Fai Wong and Tulika Mitra, *National University of Singapore*

MPK: A Compiler and Runtime for Mega-Kernelizing Tensor Programs 1909
Xinhao Cheng, Zhihao Zhang, Yu Zhou, and Jianan Ji, *Carnegie Mellon University*; Jincheng Jiang, *Tsinghua University*;
Zepeng Zhao and Ziruo Xiao, *Carnegie Mellon University*; Zihao Ye and Yingyi Huang, *NVIDIA*; Ruihang Lai, Hongyi Jin, Bohan Hou, Mengdi Wu, Yixin Dong, and Anthony Yip, *Carnegie Mellon University*; Zihao Ye, *University of Michigan*; Songting Wang, *Carnegie Mellon University*; Wenqin Yang, *Independent Researcher*; Xupeng Miao, *Peking University*; Tianqi Chen, *Carnegie Mellon University and NVIDIA*; Zhihao Jia, *Carnegie Mellon University*

GraCE: Unlocking CUDA Graphs with Compiler Support for ML Workloads 1927
Abhishek Ghosh and Ajay Nayak, *Indian Institute of Science*; Ashish Panwar, *Microsoft Research*; Arkaprava Basu, *Indian Institute of Science*

VTC: DNN Compilation with Virtual Tensors for Data Movement Elimination 1949
Muyan Hu, Ahan Gupta, and Jiachen Yuan, *University of Illinois Urbana-Champaign*; Vima Gupta, *Georgia Institute of Technology*; Taeksang Kim and Xin Xu, *University of Illinois Urbana-Champaign*; Janardhan Kulkarni and Ofer Dekel, *Microsoft*; Vikram Adve and Charith Mendis, *University of Illinois Urbana-Champaign*

Serverless Plus Resilience

Stop Pretending to be Busy: A Case for Serverless Paradigms in Co-located Batch Workloads

(Operational Systems)..... 1967

Xiaohu Chai, *Tsinghua University and Ant Group*; Jianfeng Tan, Congsi Yuan, Bowen Yang, Hao Dai, Tongkai Yang, and Chao Huang, *Ant Group*; Dong Du, *Shanghai Jiao Tong University*; Yu Chen, *Quan Cheng Laboratory and Tsinghua University*

Continuation-Centric Computing with Arca..... 1989

Akshay Srivatsan, Yuhan Deng, Katherine Mohr, Emma Sudo, Sebastian Ingino, Francis Chua, and Keith Winstein, *Stanford University*

Rethinking Process Snapshots for Near-Warm Serverless Cold Starts..... 2007

Ben Holmes, Baltasar Dinis, Lana Honcharuk, and Adam Belay, *MIT CSAIL*; Joshua Fried, *University of Pennsylvania*

Distributed Speculative Execution for Resilient Cloud Applications..... 2027

Tianyu Li, *MIT CSAIL*; Badrish Chandramouli and Philip A. Bernstein, *Microsoft Research*; Sam Madden, *MIT CSAIL*

TrainMover: An Interruption-Resilient Runtime for ML Training..... 2047

ChonLam Lao, *Harvard University and Alibaba Group*; Jiaqi Gao, Jiamin Cao, Zhipeng Zhang, Pengcheng Zhang, Jiangfei Duan, Zhilong Zheng, Yu Guan, Yichi Xu, Yong Li, and Zhengping Qian, *Alibaba Group*; Aditya Akella, *The University of Texas at Austin*; Minlan Yu, *Harvard University*; Ennan Zhai, Dennis Cai, and Jingren Zhou, *Alibaba Group*

Accelerator and Device Virtualization

MOONBRIGHT: A GPU Memory Allocator with Device-Side Page Table Materialization and

Deferred TLB Coherence..... 2065

Yangyu Zhang, *SKLP, Institute of Computing Technology, Chinese Academy of Sciences and University of Chinese Academy of Sciences*; Lei Chen, *University of Chinese Academy of Sciences*; Chunwei Xia, *University of Leeds*; Shuaijiang Li, Shuoming Zhang, Zhicheng Li, Qianqi Sun, Jiawei Xiao, Ruiyuan Xu, Ao Chen, Guangli Li, Xiaobing Feng, Huimin Cui, Chenxi Wang, and Jiacheng Zhao, *SKLP, Institute of Computing Technology, Chinese Academy of Sciences and University of Chinese Academy of Sciences*

Nixie: Efficient, Transparent Temporal Multiplexing for Consumer GPUs..... 2085

Yechen Xu, Yifei Wang, Nathanael Ren, Yiran Chen, and Danyang Zhuo, *Duke University*

μ Shell: A Microkernel-based FPGA Shell Architecture..... 2103

Jiyang Chen, Anubhav Panda, and Harshavardhan Unnibhavi, *Technical University of Munich*; Atsushi Koshiba, *Tokyo University of Science*; Pramod Bhatotia, *Technical University of Munich*

Virtualizing eBPF with Late-Binding..... 2127

Jing Zhang, *Shanghai Jiao Tong University*; Xiaguannan Song, *Harbin Institute of Technology, Shenzhen*; Dong Du, Yubin Xia, Binyu Zang, and Haibo Chen, *Shanghai Jiao Tong University*

Fleet and Cluster Scheduling

DVLA: Dynamic VM Lifetime Aware Scheduling for Drifting Lifetime Distributions and Long-Lived

VM Placement Debt (Operational Systems)..... 2149

Zhengdong Zhang, Zihan Xu, Zhidong Hu, Yanbo Shan, Fei Peng, Suhong Chen, Kaiyuan Shen, Xiangyun Kong, Handu Ding, Bing He, and Binda Ma, *Alibaba Cloud Computing*

PIMS: Fleet-wide Datacenter Maintenance with Minimal Capacity Buffer and Predictable Latency

(Operational Systems)..... 2169

Benjamin Leonhardi, *Meta Platforms*; Evangelia Kalyvianaki, *University of Cambridge*; Yang Wang, *Meta Platforms and The Ohio State University*; Abdelrahman Adam, Agshin Nabiye, Aleks Shirokov, Amitav Mohanty, Daniil Balenko, Elaine Zhao, and Essam Ewaisha, *Meta Platforms*; Hongbo Dong, *NexGeMM LLC*; Igor Marnat, Lev Novikov, and Min Zeng, *Meta Platforms*; Steven Shingler, *Independent Researcher*; Timofey Durakov, Wiliam de Abreu Pinho, Ben Christensen, Mayank Pundir, and Kaushik Veeraraghavan, *Meta Platforms*

Heterogeneity at Hyperscale: Characterization and Scheduling of Large Production AI Clusters at Alibaba (Operational Systems).....	2187
Suyi Li, <i>Hong Kong University of Science and Technology</i> ; Lingyun Yang, <i>Hong Kong University of Science and Technology and Alibaba Group</i> ; Haoxuan Yu, Sheng Yao, Tianyuan Wu, Xiaoxiao Jiang, and Hanfeng Lu, <i>Hong Kong University of Science and Technology</i> ; Kangjin Wang, <i>Alibaba Group</i> ; Chenhao Wang, <i>Fudan University</i> ; Shenglin Xu, Lun Wang, Qingyang Duan, Shenghao Liang, Xiu Lin, Meng Zhang, Wenchao Wu, Yinghao Yu, Guodong Yang, and Liping Zhang, <i>Alibaba Group</i> ; Wei Wang, <i>Hong Kong University of Science and Technology</i>	
MIMESYS: Generating Realistic Executable Testing Environments from Resource Usage Traces	2205
Donghyun Kim, Zichao Hu, Joydeep Biswas, Aditya Akella, and Daehyeok Kim, <i>The University of Texas at Austin</i>	
A Coherent, Consistent Session on Caching	
MERLIN: An Efficient Adaptive Cache Eviction Algorithm via Fine-Grained Characterization	2223
Liuja Li, Jinhao Guo, Yi Fan, and Jianyu Wu, <i>Peking University</i> ; Zhenlin Wang, <i>Michigan Tech</i> ; Jie Zhang, <i>Peking University</i> ; Yuval Tamir, <i>University of California, Los Angeles</i> ; Xiaolin Wang, Yingwei Luo, and Diyu Zhou, <i>Peking University</i>	
Learning-Augmented Heuristics: Simple yet Smart, Robust and Interpretable Cache Eviction.....	2241
Haocheng Xia, <i>Harvard University and University of Illinois Urbana–Champaign</i> ; William Nixon, <i>University of Chicago and Harvard University</i> ; Bintang Dwi Marthen, <i>Harvard University and Institut Teknologi Bandung</i> ; Pranav Bhandari, <i>Meta</i> ; Juncheng Yang, <i>Harvard University</i>	
WriteGuards: Distributed Storage Support for Strongly Consistent Caches.....	2261
Ziming Mao, <i>University of California, Berkeley</i> ; Atul Adya and Jonathan Ellithorpe, <i>Databricks</i> ; Rishabh Iyer and Matei Zaharia, <i>University of California, Berkeley</i> ; Scott Shenker, <i>University of California, Berkeley, and ICSI</i> ; Ion Stoica, <i>University of California, Berkeley</i>	
MEGALON: Efficient Data Sharing for Partly Coherent CXL Memory	2281
Jiyu Hu, Seokjoo Cho, Landon Johnson, Kiran Hombal, and Shreesha Gopalakrishna Bhat, <i>University of Illinois Urbana–Champaign</i> ; Marcos K. Aguilera, <i>NVIDIA</i> ; Ramnathan Alagappan and Aishwarya Ganesan, <i>University of Illinois Urbana–Champaign</i>	
Mobile and Edge Systems	
Inference in the Shadows: Taming Memory Bandwidth Contention in Mobile LLM Inference with Sereno	2299
Tong Xin, Xinrui Shi, Mingkai Dong, and Zeyu Mi, <i>Institute of Parallel and Distributed Systems, Shanghai Jiao Tong University</i>	
LifeLine: An Object-Page Lifetime Alignment GC Enabling Minimal Memory Copying for Mobile Devices	2319
Jiacheng Huang and Yunmo Zhang, <i>City University of Hong Kong</i> ; Qingan Li, <i>Wuhan University</i> ; Junqiao Qiu, <i>City University of Hong Kong</i> ; Chun Jason Xue, <i>Mohamed bin Zayed University of Artificial Intelligence</i>	
Unleash All Cores: Asymmetry-aware Scalable DNN Inference on Mobile CPUs	2337
Qianlong Sang, Puyi He, Huanghuang Liang, and Yili Gong, <i>Wuhan University</i> ; Chuang Hu and Xiaobo Zhou, <i>University of Macau</i> ; Dazhao Cheng, <i>Wuhan University</i>	
Surviving the Impossible Trinity: Revisiting CPU Scheduling Problem on Modern COTS Mobile Devices (Operational Systems).....	2353
Jun Xiao, Qinhui Gu, Ligeng Chen, Lizhi Sun, Zicheng Wang, Yinggang Guo, and Lu Liu, <i>Honor Device Co., Ltd.</i> ; Hao Wu, <i>Nanjing University</i> ; Borui Li, <i>Southeast University</i>	
Wild Cards	
qTPU: Hybrid Tensor Networks for Quantum-Classical Acceleration.....	2367
Nathaniel Tornow, Emmanouil Giortamis, Dennis Sprokholt, Christian Mendl, and Pramod Bhatotia, <i>Technical University of Munich</i>	
Acumen: A Platform for Encrypted and Accountable Collaborative Editing	2389
Ryan Cottone, <i>Stanford University</i> ; Darya Kaviani, Conor Power, Will Giorza, Evelyn Koo, Natacha Crooks, and Raluca Popa, <i>University of California, Berkeley</i>	
Drs.NAS: Ultra-Efficient Neural Architecture Search for Recommendation Systems	2407
Ruixuan Wang and Xun Jiao, <i>Villanova University</i>	

SMARTTalk: Teaching SMART Logs to Talk to LLMs 2423
Mayur Akewar and Dongsheng Luo, *Florida International University*; Sandeep Madireddy, *Argonne National Laboratory*;
Janki Bhimani, *Florida International University*

Tooling Potpourri

Svalinn: Overload Control in Large-Scale Servers with Multiple Resource Bottlenecks 2443
Bhaskar Subhash Pardeshi, Peidi Song, and Ahmed Saeed, *Georgia Institute of Technology*

PeeR: First-Class Scheduling for Latency-Critical eBPF Applications 2465
Jeremy Carin, Ben Holmes, and Weiyang Wang, *MIT CSAIL*; Ankit Bhardwaj, *Tufts University*; Manya Ghobadi,
MIT CSAIL and Systalyze

TypeCraft: A Lightweight Data Type Profiler with High Resolution 2483
Zecheng Li, *North Carolina State University*; Xu Liu, Namhyung Kim, Blake Jones, and Alexey Alexandrov, *Google*;
Jiajia Li, *North Carolina State University*

All Along the Watchtower: Achieving the Trinity of Observability in Cloud with DiTing 2501
Zhenyu Ren and Shuzhi Feng, *Alibaba Group*; Erci Xu, *Shanghai Jiaotong University*; Changsheng Niu, Haoyu Mao,
Beibei Wang, Chong Gao, Zhenshan Zhang, Xinrui Yu, Jiangwei Huang, Jiesheng Wu, and Hong Tang, *Alibaba Group*

The Session About Correctness

jwmalloc: A Verified Memory Allocator for Mobile Devices 2515
Jiawei Wang, Ming Fu, Ruixian Wang, and Chao Xu, *Huawei Central Software Institute*; Jonas Oberhauser, *Huawei
Central Software Institute and Huawei Hilbert Research Center*; Haibo Chen, *Huawei Central Software Institute and
Shanghai Jiao Tong University*

Neuro-Symbolic Proof Generation for Scaling Systems Software Verification 2533
Baoding He, *Nanjing University*; Zenan Li, *ETH Zurich*; Wei Sun and Yuan Yao, *Nanjing University*; Taolue Chen,
Birkbeck, University of London; Xiaoxing Ma, *Nanjing University*; Zhendong Su, *ETH Zurich*

Spain: Succinct proofs for numerical computations 2551
Zachary DeStefano, Noah Golub, Zile Huang, Julius Zhang, Sam Frank, and Michael Walfish, *NYU*

RT: Regular Types for the Streaming Shell 2583
Zekai Li, Lukas Lazarek, Evangelos Lamprou, and George Kapetanakis, *Brown University*; Konstantinos Mamouras,
Rice University; Nikos Vasilakis, *Brown University*