

2024 USENIX Annual Technical Conference

July 10–12, 2024
Santa Clara, CA, USA

Wednesday, July 10

Cloud Computing

Harmonizing Efficiency and Practicability: Optimizing Resource Utilization in Serverless Computing with JIAGU1

Qingyuan Liu, Yanning Yang, Dong Du, and Yubin Xia, *Institute of Parallel and Distributed Systems, SEIEE, Shanghai Jiao Tong University; Engineering Research Center for Domain-specific Operating Systems, Ministry of Education;* Ping Zhang and Jia Feng, *Huawei Cloud;* James R. Larus, *EPFL;* Haibo Chen, *Institute of Parallel and Distributed Systems, SEIEE, Shanghai Jiao Tong University; Engineering Research Center for Domain-specific Operating Systems, Ministry of Education; Key Laboratory of System Software (Chinese Academy of Science)*

ALPS: An Adaptive Learning, Priority OS Scheduler for Serverless Functions 19

Yuqi Fu, *University of Virginia;* Ruizhe Shi, *George Mason University;* Haoliang Wang, *Adobe Research;* Songqing Chen, *George Mason University;* Yue Cheng, *University of Virginia*

Starburst: A Cost-aware Scheduler for Hybrid Cloud 37

Michael Luo, Siyuan Zhuang, Suryaprakash Vengadesan, and Romil Bhardwaj, *UC Berkeley;* Justin Chang, *UC Santa Barbara;* Eric Friedman, Scott Shenker, and Ion Stoica, *UC Berkeley*

StreamBox: A Lightweight GPU SandBox for Serverless Inference Workflow 59

Hao Wu, Yue Yu, and Junxiao Deng, *Huazhong University of Science and Technology;* Shadi Ibrahim, *Inria;* Song Wu and Hao Fan, *Huazhong University of Science and Technology and Jinyinhu Laboratory;* Ziyue Cheng, *Huazhong University of Science and Technology;* Hai Jin, *Huazhong University of Science and Technology and Jinyinhu Laboratory*

ML Inference

Power-aware Deep Learning Model Serving with μ -Serve 75

Haoran Qiu, Weichao Mao, Archit Patke, and Shengkun Cui, *University of Illinois Urbana-Champaign;* Saurabh Jha, Chen Wang, and Hubertus Franke, *IBM Research;* Zbigniew Kalbarczyk, Tamer Başar, and Ravishankar K. Iyer, *University of Illinois Urbana-Champaign*

Fast Inference for Probabilistic Graphical Models 95

Jiantong Jiang, *The University of Western Australia;* Zeyi Wen, *HKUST (Guangzhou) and HKUST;* Atif Mansoor and Ajmal Mian, *The University of Western Australia*

Cost-Efficient Large Language Model Serving for Multi-turn Conversations with CachedAttention 111

Bin Gao, *National University of Singapore;* Zhuomin He, *Shanghai Jiaotong University;* Puru Sharma, Qingxuan Kang, and Djordje Jevdjic, *National University of Singapore;* Junbo Deng, Xingkun Yang, Zhou Yu, and Pengfei Zuo, *Huawei Cloud*

PUZZLE: Efficiently Aligning Large Language Models through Light-Weight Context Switch 127

Kinman Lei, Yuyang Jin, Mingshu Zhai, Kezhao Huang, Haoxing Ye, and Jidong Zhai, *Tsinghua University*

Storage 1

ScalaAFA: Constructing User-Space All-Flash Array Engine with Holistic Designs141

Shushu Yi, *Peking University and Zhongguancun Laboratory;* Xiurui Pan, *Peking University;* Qiao Li, *Xiamen University;* Qiang Li, *Alibaba;* Chenxi Wang, *University of Chinese Academy of Sciences;* Bo Mao, *Xiamen University;* Myoungsoo Jung, *KAIST and Panmnesia;* Jie Zhang, *Peking University and Zhongguancun Laboratory*

FASTCOMMIT: resource-efficient, performant and cost-effective file system journaling 157

Harshad Shirwadkar, Saurabh Kadekodi, and Theodore Tso, *Google*

ZMS: Zone Abstraction for Mobile Flash Storage173

Joo-Young Hwang, Seokhwan Kim, Daejun Park, Yong-Gil Song, Junyoung Han, Seunghyun Choi, and Sangyeun Cho, *Samsung Electronics;* Youjip Won, *Korea Advanced Institute of Science and Technology*

Ethane: An Asymmetric File System for Disaggregated Persistent Memory 191

Miao Cai, *College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics;* Junru Shen, *College of Computer Science and Software Engineering, Hohai University;* Baoliu Ye, *State Key Laboratory for Novel Software Technology, Nanjing University*

Networks 1

PeRF: Preemption-enabled RDMA Framework 209
Sugi Lee and Mingyu Choi, *Acryl Inc.*; Ikjun Yeom, *Acryl Inc. and Sungkyunkwan University*; Younghoon Kim, *Sungkyunkwan University*

CyberStar: Simple, Elastic and Cost-Effective Network Functions Management in Cloud Network at Scale 227
Tingting Xu, *Nanjing University*; Bengbeng Xue, Yang Song, Xiaomin Wu, Xiaoxin Peng, and Yilong Lyu, *Alibaba Group*; Xiaoliang Wang, Chen Tian, Baoliu Ye, and Camtu Nguyen, *Nanjing University*; Biao Lyu and Rong Wen, *Alibaba Group*; Zhigang Zong, *Alibaba Group and Zhejiang University*; Shunmin Zhu, *Alibaba Group and Tsinghua University*

OSMOSIS: Enabling Multi-Tenancy in Datacenter SmartNICs 247
Mikhail Khalilov, Marcin Chrapek, Siyuan Shen, Alessandro Vezzu, Thomas Benz, Salvatore Di Girolamo, and Timo Schneider, *ETH Zürich*; Daniele De Sensi, *ETH Zürich and Sapienza University of Rome*; Luca Benini and Torsten Hoefler, *ETH Zürich*

ETC: An Elastic Transmission Control Using End-to-End Available Bandwidth Perception 265
Feixue Han, *Tsinghua Shenzhen International Graduate School and Peng Cheng Laboratory*; Qing Li, *Peng Cheng Laboratory*; Peng Zhang, *Tencent*; Gareth Tyson, *Hong Kong University*; Yong Jiang, *Tsinghua Shenzhen International Graduate School and Peng Cheng Laboratory*; Mingwei Xu, *Tsinghua University*; Yulong Lan and ZhiCheng Li, *Tencent*

Edge Computing

More is Different: Prototyping and Analyzing a New Form of Edge Server with Massive Mobile SoCs 285
Li Zhang, *Beijing University of Posts and Telecommunications*; Zhe Fu, *Tsinghua University*; Boqing Shi and Xiang Li, *Beijing University of Posts and Telecommunications*; Rujin Lai and Chenyang Yang, *vclusters*; Ao Zhou, Xiao Ma, Shangguang Wang, and Mengwei Xu, *Beijing University of Posts and Telecommunications*

HiP4-UPF: Towards High-Performance Comprehensive 5G User Plane Function on P4 Programmable Switches . . . 303
Zhixin Wen and Guanhua Yan, *Binghamton University*

KEPC-Push: A Knowledge-Enhanced Proactive Content Push Strategy for Edge-Assisted Video Feed Streaming . . . 321
Ziwen Ye, *Peng Cheng Laboratory and Tsinghua Shenzhen International Graduate School*; Qing Li, *Peng Cheng Laboratory*; Chunyu Qiao, *ByteDance*; Xiaoteng Ma, *Tsinghua Shenzhen International Graduate School*; Yong Jiang, *Peng Cheng Laboratory and Tsinghua Shenzhen International Graduate School*; Qian Ma and Shengbin Meng, *ByteDance*; Zhenhui Yuan, *University of Warwick*; Zili Meng, *HKUST*

High-density Mobile Cloud Gaming on Edge SoC Clusters 339
Li Zhang, Shangguang Wang, and Mengwei Xu, *Beijing University of Posts and Telecommunications*

Operating Systems 1

Limitations and Opportunities of Modern Hardware Isolation Mechanisms 349
Xiangdong Chen and Zhaofeng Li, *University of Utah*; Tirth Jain, *Maya Labs*; Vikram Narayanan and Anton Burtsev, *University of Utah*

FetchBPF: Customizable Prefetching Policies in Linux with eBPF 369
Xuechun Cao, Shaurya Patel, and Soo Yee Lim, *University of British Columbia*; Xueyuan Han, *Wake Forest University*; Thomas Pasquier, *University of British Columbia*

Fast (Trapless) Kernel Probes Everywhere 379
Jinghao Jia, *University of Illinois Urbana-Champaign*; Michael V. Le and Salman Ahmed, *IBM T.J. Watson Research Center*; Dan Williams, *Virginia Tech and IBM T.J. Watson Research Center*; Hani Jamjoom, *IBM T.J. Watson Research Center*; Tianyin Xu, *University of Illinois at Urbana-Champaign*

HydraRPC: RPC in the CXL Era 387
Teng Ma, *Alibaba Group*; Zheng Liu, *Zhejiang University and Alibaba Group*; Chengkun Wei, *Zhejiang University*; Jialiang Huang, *Alibaba Group and Tsinghua University*; Youwei Zhuo, *Alibaba Group and Peking University*; Haoyu Li, *Zhejiang University*; Ning Zhang, Yijin Guan, and Dimin Niu, *Alibaba Group*; Mingxing Zhang, *Tsinghua University*; Tao Ma, *Alibaba Group*

EXTMEM: Enabling Application-Aware Virtual Memory Management for Data-Intensive Applications 397
Sepehr Jalalian, Shaurya Patel, Milad Rezaei Hajidehi, Margo Seltzer, and Alexandra Fedorova, *University of British Columbia*

Thursday, July 11

Operating Systems 2

- Telescope: Telemetry for Gargantuan Memory Footprint Applications** 409
Alan Nair, Sandeep Kumar, and Aravinda Prasad, *Intel Labs*; Ying Huang, *Intel Corporation*; Andy Rudoff and Sreenivas Subramoney, *Intel Labs*
- An Empirical Study of Rust-for-Linux: The Success, Dissatisfaction, and Compromise** 425
Hongyu Li, *Beijing University of Posts and Telecommunications*; Liwei Guo, *University of Electronic Science and Technology of China*; Yexuan Yang, Shangguang Wang, and Mengwei Xu, *Beijing University of Posts and Telecommunications*
- Scalable and Effective Page-table and TLB management on NUMA Systems** 445
Bin Gao, Qingxuan Kang, and Hao-Wei Tee, *National University of Singapore*; Kyle Timothy Ng Chu, *Horizon Quantum Computing*; Alireza Sanaee, *Queen Mary University of London*; Djordje Jevdjic, *National University of Singapore*
- UniMem: Redesigning Disaggregated Memory within A Unified Local-Remote Memory Hierarchy** 463
Yijie Zhong, Minqiang Zhou, and Zhirong Shen, *Xiamen University*; Jiwu Shu, *Xiamen University, and Minjiang University*

Correctness

- WINGFUZZ: Implementing Continuous Fuzzing for DBMSs** 479
Jie Liang, Zhiyong Wu, and Jingzhou Fu, *Tsinghua University*; Yiyuan Bai and Qiang Zhang, *Shuimu Yulin Technology Co., Ltd.*; Yu Jiang, *Tsinghua University*
- Balancing Analysis Time and Bug Detection: Daily Development-friendly Bug Detection in Linux** 493
Keita Suzuki, *Keio University*; Kenta Ishiguro, *Hosei University*; Kenji Kono, *Keio University*
- Kivi: Verification for Cluster Management** 509
Bingzhe Liu and Gangmuk Lim, *UIUC*; Ryan Beckett, *Microsoft*; P. Brighten Godfrey, *UIUC and Broadcom*
- MONARCH: A Fuzzing Framework for Distributed File Systems** 529
Tao Lyu, *EPFL*; Liyi Zhang, *University of Waterloo*; Zhiyao Feng, Yueyang Pan, and Yujie Ren, *EPFL*; Meng Xu, *University of Waterloo*; Mathias Payer and Sanidhya Kashyap, *EPFL*

ML Training

- Accelerating the Training of Large Language Models using Efficient Activation Rematerialization and Optimal Hybrid Parallelism** 545
Tailing Yuan, Yulian Liu, Xucheng Ye, Shenglong Zhang, Jianchao Tan, Bin Chen, Chengru Song, and Di Zhang, *Kuaishou Technology*
- Metis: Fast Automatic Distributed Training on Heterogeneous GPUs** 563
Taegeon Um, Byungsoo Oh, Minyoung Kang, Woo-Yeon Lee, Goeun Kim, Dongseob Kim, Youngtaek Kim, and Mohd Muzzammil, *Samsung Research*; Myeongjae Jeon, *UNIST*
- FwdLLM: Efficient Federated Finetuning of Large Language Models with Perturbed Inferences** 579
Mengwei Xu, Dongqi Cai, Yaozong Wu, Xiang Li, and Shangguang Wang, *Beijing University of Posts and Telecommunications (BUPT)*

Security 1

- A Secure, Fast, and Resource-Efficient Serverless Platform with Function REWIND** 597
Jaehyun Song and Bumsuk Kim, *Sungkyunkwan University*; Minwoo Kwak, *Yonsei University*; Byoungyoung Lee, *Seoul National University*; Euisong Seo, *Sungkyunkwan University*; Jinkyu Jeong, *Yonsei University*
- SimEnc: A High-Performance Similarity-Preserving Encryption Approach for Deduplication of Encrypted Docker Images** 615
Tong Sun and Bowen Jiang, *Zhejiang University*; Borui Li, *Southeast University*; Jiamei Lv, Yi Gao, and Wei Dong, *Zhejiang University*
- mmTLS: Scaling the Performance of Encrypted Network Traffic Inspection** 631
Junghan Yoon, *Seoul National University*; Seunghyun Do and Duckwoo Kim, *KAIST*; Taejoong Chung, *Virginia Tech*; KyoungSoo Park, *Seoul National University*

ML-System Co-Design

Pecan: Cost-Efficient ML Data Preprocessing with Automatic Transformation Ordering and Hybrid Placement ... 649

Dan Graur, Oto Mraz, Muyu Li, and Sepehr Pourghannad, *ETH Zurich*; Chandramohan A. Thekkath, *Google*;

Ana Klimovic, *ETH Zurich*

OPER: Optimality-Guided Embedding Table Parallelization for Large-scale Recommendation Model 667

Zheng Wang, *University of California, San Diego*; Yuke Wang, Boyuan Feng, and Guyue Huang, *University of California, Santa Barbara*; Dheevatsa Mudigere and Bharath Muthiah, *Meta*; Ang Li, *Pacific Northwest National Laboratory*;

Yufei Ding, *University of California, San Diego*

MAGPy: Compiling Eager Mode DNN Programs by Monitoring Execution States 683

Chen Zhang, Rongchao Dong, Haojie Wang, Runxin Zhong, Jike Chen, and Jidong Zhai, *Tsinghua University*

Quant-LLM: Accelerating the Serving of Large Language Models via FP6-Centric Algorithm-System

Co-Design on Modern GPUs 699

Haojun Xia, *University of Sydney*; Zhen Zheng and Xiaoxia Wu, *Microsoft*; Shiyang Chen, *Rutgers University*;

Zhewei Yao, Stephen Youn, Arash Bakhtiari, and Michael Wyatt, *Microsoft*; Donglin Zhuang and Zhongzhu Zhou,

University of Sydney; Olatunji Ruwase, Yuxiong He, and Shuaiwen Leon Song, *Microsoft*

Networks 2

QDSR: Accelerating Layer-7 Load Balancing by Direct Server Return with QUIC 715

Ziqi Wei, *Tsinghua Shenzhen International Graduate School and Peng Cheng Laboratory*; Zhiqiang Wang, *Tencent and Peng Cheng Laboratory*; Qing Li, *Peng Cheng Laboratory*; Yuan Yang, *Tsinghua University*; Cheng Luo and Fuyu Wang, *Tencent*; Yong Jiang, *Tsinghua Shenzhen International Graduate School and Peng Cheng Laboratory*; Sijie Yang, *Tencent*; Zhenhui Yuan, *Northumbria University*

Evaluating Chiplet-based Large-Scale Interconnection Networks via Cycle-Accurate Packet-Parallel Simulation ... 731

Yinxiao Feng and Yuchen Wei, *Institute for Interdisciplinary Information Sciences, Tsinghua University*; Dong Xiang, *School of Software, Tsinghua University*; Kaisheng Ma, *Institute for Interdisciplinary Information Sciences, Tsinghua University*

Config-Snob: Tuning for the Best Configurations of Networking Protocol Stack 749

Manaf Bin-Yahya, Yifei Zhao, and Hossein Shafieirad, *Huawei Technologies Canada*; Anthony Ho, *Huawei Technologies Canada and University of Waterloo*; Shijun Yin and Fanzhao Wang, *Huawei Technologies China*; Geng Li, *Huawei Technologies Canada*

Conspirator: SmartNIC-Aided Control Plane for Distributed ML Workloads 767

Yunming Xiao, *Northwestern University*; Diman Zad Tootaghaj, Aditya Dhakal, Lianjie Cao, and Puneet Sharma, *Hewlett Packard Labs*; Aleksandar Kuzmanovic, *Northwestern University*

Memory

FBMM: Making Memory Management Extensible With Filesystems 785

Bijan Tabatabai, James Sorenson and Michael M. Swift, *University of Wisconsin–Madison*

Mangosteen: Fast Transparent Durability for Linearizable Applications using NVM 799

Sergey Egorov, Gregory Chockler, and Brijesh Dongol, *University of Surrey, UK*; Dan O’Keeffe, *Royal Holloway, University of London, UK*; Sadegh Keshavarzi, *University of Surrey, UK*

FlexMem: Adaptive Page Profiling and Migration for Tiered Memory 817

Dong Xu, *University of California, Merced*; Junhee Ryu, Jinho Baek, and Kwangsik Shin, *SK hynix*; Pengfei Su and Dong Li, *University of California, Merced*

Reliability

- SuperBench: Improving Cloud AI Infrastructure Reliability with Proactive Validation** 835
Yifan Xiong, Yuting Jiang, Ziyue Yang, and Lei Qu, *Microsoft Research*; Guoshuai Zhao, Shuguang Liu, Dong Zhong, Boris Pinzur, Jie Zhang, Yang Wang, Jithin Jose, Hossein Pourreza, Jeff Baxter, Kushal Datta, Prabhat Ram, Luke Melton, and Joe Chau, *Microsoft*; Peng Cheng, Yongqiang Xiong, and Lidong Zhou, *Microsoft Research*
- Removing Obstacles before Breaking Through the Memory Wall: A Close Look at HBM Errors in the Field** 851
Ronglong Wu, Shuyue Zhou, Jiahao Lu, Zhirong Shen, and Zikang Xu, *Xiamen Key Laboratory of Intelligent Storage and Computing, Xiamen University*; Jiwu Shu, *Xiamen Key Laboratory of Intelligent Storage and Computing, Xiamen University, and Minjiang University*; Kunlin Yang and Feilong Lin, *Huawei Technologies Co., Ltd*; Yiming Zhang, *Xiamen Key Laboratory of Intelligent Storage and Computing, Xiamen University*
- MSFRD: Mutation Similarity based SSD Failure Rating and Diagnosis for Complex and Volatile Production Environments** 869
Yuqi Zhang, Tianyi Zhang, Wenwen Hao, Shuyang Wang, Na Liu, and Xing He, *Samsung R&D Institute China Xi'an, Samsung Electronics*; Yang Zhang, Weixin Wang, Yongguang Cheng, Huan Wang, Jie Xu, Feng Wang, and Bo Jiang, *ByteDance Inc.*; Yongwong Gwon, Jongsung Na, Zoe Kim, and Geunrok Oh, *Samsung Electronics*

Friday, July 12

Deployed Systems

- Diagnosing Application-network Anomalies for Millions of IPs in Production Clouds** 885
Zhe Wang, *Shanghai Jiao Tong University*; Huanwu Hu, *Alibaba Cloud*; Linghe Kong, *Shanghai Jiao Tong University*; Xinlei Kang and Teng Ma, *Alibaba Cloud*; Qiao Xiang, *Xiamen University*; Jingxuan Li and Yang Lu, *Alibaba Cloud*; Zhuo Song, *Shanghai Jiao Tong University and Alibaba Cloud*; Peihao Yang, *Alibaba Cloud*; Jiejian Wu, *Shanghai Jiao Tong University*; Yong Yang and Tao Ma, *Alibaba Cloud*; Zheng Liu, *Alibaba Cloud and Zhejiang University*; Xianlong Zeng and Dennis Cai, *Alibaba Cloud*; Guihai Chen, *Shanghai Jiao Tong University*
- Data Caching for Enterprise-Grade Petabyte-Scale OLAP** 901
Chunxu Tang and Bin Fan, *Alluxio*; Jing Zhao and Chen Liang, *Uber, Inc.*; Yi Wang and Beinan Wang, *Alluxio*; Ziyue Qiu, *Carnegie Mellon University and Uber, Inc.*; Lu Qiu, Bowen Ding, Shouzhao Sun, Saiguang Che, Jiaming Mai, Shouwei Chen, Yu Zhu, and Jianjian Xie, *Alluxio*; Yutian (James) Sun, *Meta, Inc.*; Yao Li and Yangjun Zhang, *Uber, Inc.*; Ke Wang, *Meta, Inc.*; Mingmin Chen, *Uber, Inc.*
- Full Lifecycle Data Analysis on a Large-scale and Leadership Supercomputer: What Can We Learn from It?** 917
Bin Yang, *Tsinghua University and National Supercomputer Center in Wuxi*; Hao Wei, *Tsinghua University*; Wenhao Zhu, *Shandong University and National Supercomputer Center in Wuxi*; Yuhao Zhang, *Tsinghua University*; Weiguo Liu, *Shandong University*; Wei Xue, *Tsinghua University, Qinghai University and Intelligent Computing, and Application Laboratory of Qinghai Province and National Supercomputer Center in Wuxi*

Wide Area Network

- Panorama: Optimizing Internet-scale Users' Routes from End to End** 935
Geng Li, Shuihai Hu, and Kun Tan, *Huawei Technologies*
- Enhancing Resource Management of the World's Largest PCDN System for On-Demand Video Streaming** 951
Rui-Xiao Zhang, *UIUC*; Haiping Wang, Shu Shi, Xiaofei Pang, Yajie Peng, and Zhichen Xue, *ByteDance*; Jiangchuan Liu, *Simon Fraser University*
- TileClipper: Lightweight Selection of Regions of Interest from Videos for Traffic Surveillance** 967
Shubham Chaudhary and Aryan Taneja, *IIT Delhi*; Anjali Singh, *Indira Gandhi Delhi Technology University for Women*; Purbasha Roy, Sohun Sikdar, Mukulika Maity, and Arani Bhattacharya, *IIT Delhi*

Virtualization

Expeditious High-Concurrency MicroVM SnapStart in Persistent Memory with an Augmented Hypervisor 985
Xingguo Pang, Yanze Zhang, and Liu Liu, *University of Macau*; Dazhao Cheng, *WuHan University*; Chengzhong Xu and Xiaobo Zhou, *University of Macau*

Taming Hot Bloat Under Virtualization with HUGESCOPE 999
Chuandong Li, *National Key Laboratory for Multimedia Information Processing, School of CS, Peking University, and Zhongguancun Laboratory*; Sai Sha, *National Key Laboratory for Multimedia Information Processing, School of CS, Peking University, and Beijing Huawei Digital Technologies*; Yangqing Zeng and Xiran Yang, *National Key Laboratory for Multimedia Information Processing, School of CS, Peking University*; Yingwei Luo and Xiaolin Wang, *National Key Laboratory for Multimedia Information Processing, School of CS, Peking University, and Zhongguancun Laboratory*; Zhenlin Wang, *Michigan Tech*; Diyu Zhou, *National Key Laboratory for Multimedia Information Processing, School of CS, Peking University, and EPFL*

CrossMapping: Harmonizing Memory Consistency in Cross-ISA Binary Translation1013
Chen Gao and Xiangwei Meng, *Lanzhou University*; Wei Li, *Tsinghua University*; Jinhui Lai, *Lanzhou University*; Yiran Zhang, *Beijing University of Posts and Telecommunications*; Fengyuan Ren, *Lanzhou University and Tsinghua University*

Security 2

Efficient Decentralized Federated Singular Vector Decomposition 1029
Di Chai, Junxue Zhang, Liu Yang, and Yilun Jin, *Hong Kong University of Science and Technology*; Leye Wang, *MoE Key Lab of High Confidence Software Technologies, Peking University, and School of Computer Science, Peking University*; Kai Chen, *Hong Kong University of Science and Technology*; Qiang Yang, *Hong Kong University of Science and Technology and Webank*

Models on the Move: Towards Feasible Embedded AI for Intrusion Detection on Vehicular CAN Bus. 1049
He Xu, Di Wu, Yufeng Lu, and Jiwu Lu, *Hunan University and ExponentiAI Innovation*; Haibo Zeng, *Virginia Tech*

CPC: Flexible, Secure, and Efficient CVM Maintenance with Confidential Procedure Calls. 1065
Jiahao Chen, *Institute of Parallel and Distributed Systems, SEIEE, Shanghai Jiao Tong University; Engineering Research Center for Domain-specific Operating Systems, Ministry of Education, China; Shanghai Key Laboratory of Scalable Computing and Systems, Shanghai Jiao Tong University*; Zeyu Mi and Yubin Xia, *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; Shanghai Key Laboratory of Scalable Computing and Systems, Shanghai Jiao Tong University*

Storage 2

RL-Watchdog: A Fast and Predictable SSD Liveness Watchdog on Storage Systems. 1083
Jin Yong Ha, *Seoul National University*; Sangjin Lee, *Chung-Ang University*; Heon Young Yeom, *Seoul National University*; Yongseok Son, *Chung-Ang University*

Exploit both SMART Attributes and NAND Flash Wear Characteristics to Effectively Forecast SSD-based Storage Failures in Clusters.1101
Yunfei Gu and Chentao Wu, *Shanghai Jiao Tong University*; Xubin He, *Temple University*

StreamCache: Revisiting Page Cache for File Scanning on Fast Storage Devices1119
Zhiyue Li and Guangyan Zhang, *Tsinghua University*

Scalable Billion-point Approximate Nearest Neighbor Search Using SmartSSDs 1135
Bing Tian, Haikun Liu, Zhuohui Duan, Xiaofei Liao, Hai Jin, and Yu Zhang, *Huazhong University of Science and Technology*

Hardware

- gVulkan: Scalable GPU Pooling for Pixel-Grained Rendering in Ray Tracing**1151
Yicheng Gu, Yun Wang, Yunfan Sun, Yuxin Xiang, Yufan Jiang, Xuyan Hu, Zhengwei Qi, and Haibing Guan,
Shanghai Jiao Tong University
- vFPIO: A Virtual I/O Abstraction for FPGA-accelerated I/O Devices**1167
Jiyang Chen, Harshavardhan Unnibhavi, Atsushi Koshiba, and Pramod Bhatotia, *Technical University of Munich*
- ScalaCache: Scalable User-Space Page Cache Management with Software-Hardware Coordination** 1185
Li Peng and Yuda An, *Peking University*; You Zhou, *Huazhong University of Science and Technology*; Chenxi Wang,
University of Chinese Academy of Sciences; Qiao Li, *Xiamen University*; Chuanning Cheng, *Huawei*; Jie Zhang,
Peking University and Zhongguancun Laboratory
- Centimani: Enabling Fast AI Accelerator Selection for DNN Training with a Novel Performance Predictor** 1203
Zhen Xie, *Binghamton University*; Murali Emani, *Argonne National Laboratory*; Xiaodong Yu, *Stevens Institute of
Technology*; Dingwen Tao, *Indiana University*; Xin He, *Xidian University*; Pengfei Su, *University of California, Merced*;
Keren Zhou, *George Mason University*; Venkatram Vishwanath, *Argonne National Laboratory*

Potpourri

- A Difference World: High-performance, NVM-invariant, Software-only Intermittent Computation** 1223
Harrison Williams, *Virginia Tech*; Saim Ahmad, *Amazon*; Matthew Hicks, *Virginia Tech*
- Efficient Large Graph Processing with Chunk-Based Graph Representation Model** 1239
Rui Wang, *Zhejiang University and Hangzhou High-Tech Zone (Binjiang) Institute of Blockchain and Data Security*;
Weixu Zong, Shuibing He, Xinyu Chen, Zhenxin Li, and Zheng Dang, *Zhejiang University*
- SlimArchive: A Lightweight Architecture for Ethereum Archive Nodes** 1257
Hang Feng, Yufeng Hu, and Yinghan Kou, *Zhejiang University*; Runhuai Li and Jianfeng Zhu, *BlockSec*; Lei Wu and
Yajin Zhou, *Zhejiang University*
- Every Mapping Counts in Large Amounts: Folio Accounting** 1273
David Hildenbrand, *Technical University of Munich and Red Hat GmbH*; Martin Schulz, *Technical University of Munich*;
Nadav Amit, *Technion, Israel Institute of Technology*