2026 IEEE 46th International Conference on Distributed Computing Systems (ICDCS 2026)

Seoul, Republic of Korea
22-25 June 2026

Pages 1–740

IEEE Catalog Number: CFP26040-POD
ISBN: 979-8-3195-2980-0

Copyright © 2026, IEEE

All Rights Reserved

Copyright and Reprint Permissions:

Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

*** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.

IEEE Catalog Number:	CFP26040-POD
ISBN (Print-On-Demand):	979-8-3195-2980-0
ISBN (Online):	979-8-3195-2979-4
ISSN:	1063-6927

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA

Phone:	(845) 758-0400
Fax:	(845) 758-2633
E-mail:	curran@proceedings.com
Web:	www.proceedings.com



TABLE OF CONTENTS

ZeroSwap: A Practical Solution to the Double Swapping Problem in Virtualized Environments 1
Luc Mahop, Kilian Kemgne, Baptiste Lepers, Fabienne Boyer, Alain Tchana

Sandwich: Synergizing Hierarchical Coordination with Fine-Grained Serverless Orchestration 12
Yunfeng Zhao, Chao Qiu, Weijie Guo, Chenxuan Hou, Yuting Li, Xiaofei Wang, Qinghua Hu

CARL: Crosstalk-Aware Quantum Compilation Based on Reinforcement Learning 23
Xiangzhi Zhang, Yu Liu, Yingling Mao, Xu Xu, Yuanyuan Yang

Providing Cross-Chain Interoperability Through Decentralized Message Routing and Delivery 34
Amin Rezaei, Solomon L. Davidson, Bernard Wong

Semi-Parallel Proof-Of-Work: a Practical Protocol with Improved Incentive Compatibility 45
Mustafa Doger, Sennur Ulukus

From Alpha to Omega: Lifecycle-Aware Forgetting Defense in Federated Continual Learning for Planetary Exploration 56
Beining Wu, Jun Huang, Yanxiao Zhao

MixServe: An Automatic Distributed Serving System for MoE Models with Hybrid Parallelism Based on Fused Communication Algorithm 67
Bowen Zhou, Jinrui Jia, Wenhao He, Yong Zhang, Fang Dong

Tackling the Runtime Context Fluctuation: Steady Scheduling in Data Stream Analytics for Industrial Internet of Things 78
Shuyu Cao, Wenhui Zhou, Jingyi Ning, Lei Xie

Graph Attention-Based Cooperative Multi-Agent Reinforcement Learning for Low-Latency Reliable Task Replication in Mobile Edge Computing 89
Anam Nasir, Xiang He, Teng Wang, Haomai Shi, Zhongjie Wang

PARS: Optimizing High-Dimensional Vector Search with Dimensional Reduction and Ray Tracing Core 100
Yiling Ma, Mengbai Xiao, Zhixiong Xiao, Yuan Yuan, Dongxiao Yu, Guanghui Zhang, Feng Li

Optimizing Wide-Area Collective Communication via Flow-Dependency-Aware Rate Allocation 111
Yining Wang, Toon Craps, Baochun Li

FractalGPU: Fair and Elastic GPU Sharing for General-Purpose Computing 122
Jingyi Chen, Kaicheng Guo, Lingyun Yang, Yun Wang, Wenda Tang, Dongdong Ma, Pengwei Du, Qian Da Zhengwei Qi

Turmeric: A Fault-Tolerant Publish/Subscribe Service with Multiple Ordering Guarantees 133
Luís Chula, Paulo Matos, Pedro Mascarenhas, Hervé Paulino

GMACAuth: Generative Continuous Authentication via Memory-Augmented Causal Attention 144
Yantao Li, Yuhang Zhou, Hailong Hu, Huafeng Qin

QCLink: Offloading Hybrid Quantum-Classical Computation to SmartNICs via RDMA 155
Xingdong Li, Yongzhuo Lu, Zhipeng Liu, Xiaofei Yue, Junyu Chen, Yu Peng, Tingting Li, Zhaoxuan Li Ziming Zhao, Jianwei Yin

Sketch-Based in-Network Compression for Iot Satellite Networks 166
Ting-Yuan Wen, Kate Ching-Ju Lin

Competitive Caching for Distributed Data Access 177
Tianyu Zuo, Xueyan Tang, Bu Sung Lee

Competitive Transaction Admission in PCNs: Online Knapsack with Positive and Negative Items 188
Marcin Bienkowski, Julien Dallot, Dominik Danelski, Maciej Pacut, Stefan Schmid

Hash-Based Asynchronous MVBA with Optimal Complexity and Near-Optimal Resilience 197
Jovan Komatovic, Joachim Neu, Tim Roughgarden

SPML: A Proactive Congestion Control Scheme for Periodic Flows in Distributed Systems 208
Shengwen Zhou, Jing Shao, Jiawei Huang

Keep Fresh Digital Twins in UAV-Assisted IoT Networks by Exploiting Data Correlations 219
Qunli Shen, Jing Li, Jian Peng, Zichuan Xu, Cheng Zhao, Pan Zhou, Weifa Liang, Xiaohua Jia, Sajal K. Das Wenzheng Xu

OrbitBFT: Enabling Scalable and Robust BFT Consensus in LEO Constellations 230
Tianyi Sun, Shuo Liu, Minghui Xu, Xiuzhen Cheng

BeamGes: Motion Resilient Hand Gesture Recognition on Smartphones Using Beamforming Feedback Information 241
Zijie Tang, Zejun Xu, Yan Wang, Yingying Chen

Accelerating Cold Starts of On-Device LLMs via Multi-Source Inference-Aware Parameter Loading 252
Shucheng Li, Zhenfeng Wang, Fei Zeng, Zhanxi Li, Peng Yang, Feng Lyu

HandoffFL: Efficient Communication in Federated Learning Via Asynchronous Layer-Wise Handoff Training 263
Wenjun Zhang, Xiaoli Liu, Sasu Tarkoma

EA-ONS: A Deterministic Time Synchronization Framework for 5G–TSN Integrated LEO Constellations 274
Siqi Yang, Zonghui Li, Chaoqun You, Yue Gao

ShadowClone: Accelerating Cross-Shard Transactions via Shadow Accounts 285
Jiahao Qi, Dian Ding, Feilong Lin, Haotian Zhu, Jie Li, Shengyun Liu, Guangtao Xue, Jiannong Cao

Justitia-L: Budget-Constrained Fairness Optimization in Sharded Blockchains via Lagrangian Dual Control 296
Sihua Wang, Jian Zheng, Huawei Huang

Aiming Low-Latency Atomicity for Cross-Shard Transactions in a Sharding Blockchain 306
Huawei Huang, Jian Zheng, Qinde Chen, Yuting Yang

Folium: Decoupling Transaction Execution from Consensus via Follower Nodes in a Blockchain 317
Xiaofei Luo, Huawei Huang, Jian Zheng, Baozhou Xie

Clearing MCP Navigation Fog with Economics-Aware Hierarchical Tool Routing 327
Jieling Yu, Zhenheng Tang, Ruiting Zhou, Baochun Li, Bo Li

TurboInfer: Targeting Age of Model Inference Optimization for Joint Model Inference in Edge Cloud Systems 338
Chenxuan Hou, Chao Qiu, Tiantian Cao, Yunfeng Zhao, Chengwei Wang, Xiaofei Wang

Maestro: Workload-Aware Cross-Cluster Scheduling for LLM-based Multi-Agent Systems 349
Jinghao Wang, Xiao Zhou, Xiaoyang Sun, Yihui Zhang, Yilong Li, Tianyu Wo, Xu Wang, Chunming Hu Renyu Yang

EDAgent: A Concurrent Orchestration Method for Collaborative LLM Multi-Agent System 360
Zongyang Yuan, Zechang Zhang, Qinbin Li, Lailong Luo, Deke Guo, Mingrui Lao

Efficient KV Cache Migration for Geo-Distributed LLM Inference in Collaborative Edge Computing 371
Mingjin Zhang, Jiannong Cao, Tao Wu, Xiangchun Chen, Ne Wang

ReCache: Cost-Efficient DNN Preprocessing on Edge Servers via Group-Based Data Reuse 382
Chang Xu, Zhixiong Niu, Juan Fang, Qian Wang

KV Cache Reuse for Elastic LLM Inference on Edge Devices 393
Peishuo Wang, Zhenzhe Zheng, Xiaoyao Huang, Jie Wu, Fan Wu, Guihai Chen

On Hierarchical Caching with Delayed Hits 404
Kanghuai Liu, Xueyan Tang

Hardening Output Privacy for Secure Inference: A Lightweight Realization via Distributed Trust 415
Xinqian Wang, Xiaoning Liu, Shangqi Lai, Xun Yi, Ibrahim Khalil, Kwok-Yan Lam

PriVAR: Client-Side Privacy Framework for Real-Time Location-Based Augmented Reality 426
Shafizur Rahman Seeam, Ye Zheng, Zhengxiong Li, Yidan Hu

FedPF: Accurate Target Privacy Preserving Federated Learning Balancing Fairness and Utility 437
Kangkang Sun, Jun Wu, Minyi Guo, Jianhua Li, Jianwei Huang

Multi-Shuffler: Decentralized Shuffling for Differential Privacy in Mobile Edge Computing 448
Wenjie Chen, Xinhao Huang, Yubo Zhu, Tingxuan Han, Wei Tong, Sheng Zhong

Pulse: Training Acceleration for Large Diffusion Models with Automatic Pipeline Parallelism 459
Boran Sun, Guoyong Jiang, Lin Zhang, Chen Chen, Yuechen Tao, Zhishu Che, Jieling Yu, Shan Chang Huaxi Gu, Fangming Liu, Bo Li

FEDLight: A Fuel-Economic Reinforcement Learning Model for Distributed Traffic Signal Control with Spatiotemporal Decomposition 470
Lang Qian, Peng Sun, Jiayue Jin, Yang Liu, Azzedine Boukerche, Liang Song

Decentralized Heterogeneous Task Offloading in Edge Computing: A Mean-Field Learning Approach 480
Shuhui Chu, Chengzhong Xu

Efficient Deployment of Lightweight LLMs on Edge Devices: Kernel-Level Profiling and Cross-Platform Insights 491
Xinyao Liu, Xingzhou Zhang, Weisong Shi

Prism: Proactive Workload-Aware Optimization for Hybrid-Service in LMaaS Systems 502
Yuting Li, Chao Qiu, Shaoyuan Huang, Tengwen Zhang, Yunfeng Zhao, Xiaofei Wang

Incentivizing and Orchestrating Cloud-Edge LLM Speculative Decoding via Auctions 513
Mingtao Ji, Lei Jiao, Bin Tang, Zhihao Qu, Chen Chen, Baoliu Ye

HiAsCC: Hierarchical Asynchronous Collective Communication Method for Large Model Training 524
Zhihang Tang, Chenxi Li, Bo He, Qi Qi, Yulong Tao, Jingyu Wang, Laiping Zhao, Keqiu Li

MEGATRACE: Troubleshooting Hang and Slowdown in Large-Scale LLM Training Clusters 535
Fangzheng Jiao, Menghao Zhang, Bolin Chen, Jiaxun Huang, Yanmin Jia, Xiaohe Hu, Bohua Xu Bowen Han, Chunming Hu

Falconf: Configuration Error Diagnosis via Log Sequence Learning and Automated Misconfiguration Injections 546
Youyang Kim, Sahil Suneja, Yunja Choi, Young-Woo Kwon, Byungchul Tak

ShadowLLM: Resource-Efficient Hot Standby for Heterogeneous Edge LLM Serving 557
Zhenguo Chen, Pujun Ding, Yuanwei Zhu, Hao Zhang, Jing Lv, Yakun Huang, Xiuquan Qiao

D ${ }^{2}$ Write: Accelerating Erasure-Coded Writes With Distributed Encoding and Decoupled Transmission 568
Canghai Yang, Wanyi Guo, Zhiwang Yu, Chaoxia Qin, Kan Zhong, Duo Liu

Signet: Scalable Network-Driven Proof of Notification for Blockchain Systems 579
Elham Ehsani Moghadam, Marc Wyss, Jonghoon Kwon, Marc Frei, Yih-Chun Hu, Adrian Perrig Alberto Sonnino

Anchor Drag Attack: Exploiting Information Asymmetry in Bitcoin's Stratified Topology 590
Zihan Wang, Feiyang Chen, Yuxiang Zheng, Jiong Lou, Wugedele Bao, Celimuge Wu, Wei Zhao, Jie Li

L-SDFL: A Lightweight Semi-Distributed Federated Learning for Communication Efficiency 601
Dapeng Qu, Longtao Li, Ruiduo Li, Xiaojie Liu, Haozhe Wang, Geyong Min

Byzantine Reliable Broadcast on Partially Connected Networks Under Message Adversaries 611
Silvia Bonomi, Lorenzo Di Filippo, Sebastien Tixeuil

Synthesizing Algorithms to Avoid an Obstacle with a Swarm of Robots 622
Karine Altisen, Anaïs Durand, Pascal Lafourcade, Oussama Nahnah

Online Graph Embedding in Star Graphs 633
Julien Dallot, Darya Melnyk, Maciej Pacut, Stefan Schmid

MS-TBImap: A Real-Time Framework for Multi-Scale Temporary Boundary Integration 642
Hengyu Zhou, Weizhen Han, Xiang Shao, Haiyong Shi, Shihong Cui, Bingyi Liu

Enabling Realtime Stream Processing with Partial Computation Offloading 653
Jiamei Lv, Cheng Guo, Wenzhao Zhang, Renjie Li, Yongqi Chen, Yixiao Teng, Yi Gao, Wei Dong

Quantization-Aware Incentive Mechanism for Communication-Efficient Federated Learning 664
Hengrui Cui, Zhihao Qu, Baoliu Ye, Tao Zhuang, Bin Tang, Wenxuan Zhou

Dual-Blade: Dual-Path NVMe-Direct KV-Cache Offloading for Edge LLM Inference 675
Bodon Jeong, Hongsu Byun, Youngjae Kim, Weikuan Yu, Kyungkeun Lee, Jihoon Yang, Sungyong Park

GraphFlash: Enabling Fast and Elastic Graph Processing on Serverless Infrastructure 686
Chen Zhao, Parsa Poorsistani, Mohammad Goudarzi, Tawfiq Islam, Adel N. Toosi

RcLLM: Accelerating Generative Recommendation via Beyond-Prefix KV Caching 697
Zhan Zhao, Yuxin Wang, Amelie Chi Zhou

BCD-Megatron: A Cost-Effective Training System for Large Language Models 708
Zeyu Liu, Yan Li, Yunquan Zhang, Boyang Zhang, Guoyong Jiang, Xin Zhang, Limin Xiao, Weifeng Zhang Daning Cheng

Memory-Optimized Masked CRYSTALS-Kyber Implementation on ARM Cortex-M4 719
Ruiqi Hou, Yiwen Gao, Wei Cheng, Yuejun Liu, Jingdian Ming, Yongbin Zhou

Non-Interactive Distributed Key Management With Pre-Determined Shares 730
Yaqing Song, Yuan Zhang, Shiyu Li, Weijia Li, Zeqi Lai, Qiang Tang

TimeProtect: Time-Bound Policy Enforcement for Metadata-Hiding Encrypted Analytics 741
Qingyuan Xie, Zhao Bai, Chengjun Cai, Xiaohua Jia

CITADEL: Fortifying IoT via Context-Aware Offloading of Post-Quantum Cryptography 752
Ehan Sohn, Changjong Kim, Sunggon Kim

Spatio-Temporal Parallelism for Diffusion Model Inference on Heterogeneous Multi-GPU Systems 763
Han Liang, Jiahui Zhou, Zicheng Zhou, Xiaoxi Zhang, Xu Chen

Tensor Parallelism for State-Space Models 774
Anurag Dutt, Nimit Shah, Hazem Masarani, Anshul Gandhi

Privacy-Preserving Chunk Scheduling in a BitTorrent Implementation of Federated Learning 785
Naicheng Li, Javad Dogani, Rui Wang, Kaitai Liang, Nikolaos Laoutaris

Exploring the Feasibility and Performance of Distributed Llm Serving on Consumer-Grade Gpus 796
Lewei Jin, Yang Liu, Yongqi Chen, Xiaoqiang Guo, Kui Zhang, Yi Gao, Wei Dong

FlashServe: Adaptive Kernel Provision for Quantized LLM Serving 807
Yuyan Chen, Haodong Wang, Junyuan Liang, Wuhui Chen, Zicong Hong, Song Guo, Ruiyan Zhuang Yi Quan

Llm-Guided Training and Llm-Free Inference for Reliable Task Offloading in Edge Computing 818
Hao Guo, Kaixiang Xv, Lei Yang

Energy-Aware Usv-Uav Cooperative Task Offloading Optimization in Water Monitoring System 829
Chaogang Tang, Tiyu Yao, Shuo Xiao, Huaming Wu, Ruidong Li

Batchmon: Exploiting Serverless Functions for Cost-Effective, SLO-Driven Batch Inference Serving 840
Josep Calero-Santo, Marc Sanchez-Artigas

BalISAC-MAPPO: A MARL for Spacing Trade-Off Balancing in ISAC Multi-UAV Terrain Coverage 851
Zheyi Sun, Yang Yang, Cheng Zhan, Xianlong Jiao

AT-Cache: Towards Traffic-Aware Adaptive TCAM Rule Caching Framework 862
Lei Guo, Zeyu Luan, Qing Li, Zhuochen Fan, Bo Tang

Congestion Adaptive Load Balancing with in-Network Reordering for Datacenter Networks 873
Jiacheng Qu, Shaojun Zou, Yi Jiang

OptiVibe: Keystroke Inference Attacks Through a New Optical-Vibration Side Channel 884
Youngtak Cho, Sanket Suresh Badgujar, Srinivasan Murali, Xuhao Xie, Ming Li

A Replay-Attack Resistant and Lightweight Token Validation Framework With Hash Chain 895
Gyeongheon Jeong, Jaeho Nam, Honggeun Park, Hyunsoo Kim, Jinhwan Jung, Jongsung Kim Taekyoung Kwon

Privacy Attacks on Stable Marriage 906
Stephan A. Fahrenkrog-Petersen, Aleksander Figiel, Darya Melnyk, Tijana Milentijević, Stefan Schmid

Heterogeneous Graph Backdoor Attack 916
Jiawei Chen, Lusi Li, Daniel Takabi, Masha Sosonkina, Rui Ning

ProtoGenesis: Prototype Training with Zero Local Samples for Heterogeneous Clients in Fully Decentralized Federated Systems 927
Xianbo Wang, Shan Chang, Minghui Dai, Hao Yu, Yunnan Tu, Hongzi Zhu, Bo Li

mmLite: Leveraging mmWave Radar for Efficient Seated Pose Estimation 938
Ruili Shi, Shuai Wang, Wenchao Jiang, Zhimeng Yin, Wei Gong, Shuai Wang

RAGNAROS: Large-Scale Regex Processing Acceleration with Neural Networks for DPI 949
Sebin Jo, Taejune Park

Trilogy: Tag Information Collection in Multi-Category Commodity RFID Systems 960
Yanyan Wang, Zhihao Qu, Jia Liu, Xin He, Yingchi Mao, Bin Tang, Baoliu Ye

A Cyber-Physical Cooperative Scheme for Energy Scheduling and Dynamic Pricing in Charging Station Alliances via Cloud-Edge Intelligence 971
Bing Lin, Kai He, Xingyuan Lin, Jingtao Chen, Huaming Wu

Lucid: Lidar-Based Uav Classification and Detection for Airspace Surveillance 982
Deokjin Kim, Jeongyeup Paek, Saewoong Bahk

TriLoRa: Improving LoRa PHY Data Rate for High-Density LoRaWAN 993
Yanbo Zhang, Qirui Liu, Songfan Li, Mo Li

Rethinking Quantum Network Design Using a Verification-Based Quantum Transmission Protocol 1004
Yiming Zeng, Zhengyu Wu, Xuan Du Trinh, Yuanyuan Yang, Nengkun Yu, Aruna Balasubramanian

Resource Allocation for Surface Code Teleportation in Distillation-Based Quantum Networks 1015
Tianjie Hu, Jindi Wu, Qun Li

AURORA-Q: Asynchronous Unified Resource Optimizer for Quantum Simulation on HPC System 1026
Changjong Kim, Alex Sim, Kesheng Wu, Houjun Tang, Yongseok Son, Jisung Park, Sunggon Kim

Resource Allocation for Distributed Fault-Tolerant Quantum Computing With Surface Codes 1037
Xu Xu, Yu Liu, Yingling Mao, Yuanyuan Yang

GBSI: A GPU-Based Secondary Indexing Mechanism for Write-Intensive Workloads 1048
Yijie Zhang, Jiqing Song, Ji Li, Chuanwen Li, Shuai Tang, Yu Gu, Guodong Wang

cGraph: A Compact and Efficient Graph-Based Index for Approximate Nearest Neighbor Search 1059
Yu Liu, Mengbai Xiao, Jing Qiao, Yuan Yuan, Qi Luo, Guanghui Zhang, Dongxiao Yu

When Truth is Unverifiable: A Reputation-Based Oracle with Provable Equilibrium and Incentive-Compatible Aggregation 1070
Xiaoqin Feng, Zixuan Cheng

FedRFF: Enhanced Federated Random Fourier Feature Framework for IoT Anomaly Detection 1081
Chaoqun Li, Tao Luo, Keyuan Qiu, Jinyao Liu, Xianglong Zhang, Huanle Zhang, Si Wu, Feng Li Pengfei Hu

ARTSN: Exact and Adaptive Self-Triggered Traffic Scheduling for ARTS Networks 1092
Ruide Cao, Shuangping Zhan, Jiashuo Lin, Yan Liu, Chenxi Ling, Yi Wang, Guoming Tang

ChainCred: Enforcing Comprehensive Credential Disclosure in WEB3 Systems 1103
Rui Song, Bin Xiao, Yu Liu, Yuanyuan Yang

Integrated Load-Balanced Scheduling for Human-Vehicle Collaborative Urban Sanitation 1114
Lingzi Zhao, Huali Lu, Hao Wu, Shucheng Li, Longye Li, Tao He, Wenlong Liao, Feng Lyu

Distributed Covert Link Scheduling in Open Wireless Networks 1125
Li Yang, Yuhang Zhang, Yifei Zou, Hui Ge, Xingquan Li, Jia Yu, Dongxiao Yu

Physics-Inspired Decomposition for Weak Low-Rank Spatiotemporal Completion in Sparse Crowdsensing 1136
Mijia Zhang, En Wang, Wenbin Liu, Junwei Zhao, Yang Xu, Bo Yang, Jie Wu

BridgeLoRA: Privacy-Preserving Collaborative Skip-Layer Connectors for Efficient Transformer Fine-Tuning at the Edge 1147
Vilhelm Toivonen, Xiang Su, Xiaoli Liu, Sasu Tarkoma, Pan Hui

Minimizing Qubit Overhead in Error-Aware Distributed Quantum Computing 1158
Ranjani G. Sundaram, Himanshu Gupta

COAR: Computation-Aware Erasure-Coded Repair in Storage Clusters for Distributed Data Analytics 1169
Mingqi Li, Yuchong Hu, Lin Wang, Chenxuan Yao, Patrick P. C. Lee, Dan Feng

Deceptive Electricity Theft: New Attacks and Countermeasures in Multiple-Pricing Smart Grids 1180
Chengpeng Huang, Shang Gao, Qingqing Gan, Guyue Li, Bin Xiao

AgentDID: Trustless Identity Authentication for AI Agents 1191
Minghui Xu, Xiaoyu Liu, Yihao Guo, Chunchi Liu, Yue Zhang, Xiuzhen Cheng

Trustworthy-PWS: An Operator-Agnostic Public Warning System (PWS) Security Architecture in 5G 1202
Sourav Purification, Sang-Yoon Chang

A Robust and Secure Decentralized Handover Authentication Scheme for HetNets 1213
Wenhao Wang, Yu Li, Minze Xu, Kang Liu, Yuan Zhang, Sheng Zhong

GumSwap: Griefing-Free Universal Multi-Party Atomic Swaps 1224
Dongkun Hou, Yuanzhe Zhang, Shujie Cui, Tsz Hon Yuen, Joseph K. Liu, Jiangshan Yu

Fast Semantic Retrieval with Balanced Load and Implicit Privacy in Large-Scale Internet of Agents 1235
Jinkai Zheng, Tom H. Luan, Yuan Chang, Yuntao Wang, Zhendong Li, Haixia Peng, Xianhua Yu Nan Cheng, Zhou Su

LSCFL: Clustered Federated Learning with Label Semantics for Label-Skewed Non-IID Data 1246
Jiaxin Liu, Chunqiang Hu, Hui Xia, Meng Han, Ruinian Li, Jiguo Yu

Towards Semantic-Aware Edge-Cloud Collaboration for Cost-Efficient Video Understanding 1257
Cheng Pan, Cong Zhang, Danyang Song, Handi Chen, Edith C. H. Ngai, Jiangchuan Liu Victor C. M. Leung

FloodGuard: A Prediction-Control Closed Loop for Mitigating Cold-Start Floods in Cloud Services 1268
Jiacheng Cui, Yuwei Li, Junyu Xue, Guoming Tang

ObliMIG: Enabling Data Migration on Oblivious Storage without Interruption 1279
Bo Zhang, Helei Cui, Zhe Peng, Yu Hua, Zhiwen Yu, Bin Guo

THEIA: Scalable and Effective Data Prefetching for Geo-Distributed Erasure-Coded Storage 1290
Marc Tranzer, Shadi Ibrahim

SpecPool: Storage-Layer Speculation for Parallel Smart Contract Execution 1301
Francisco Rola, Miguel Matos, Michał Nazarewicz, Paolo Romano

ClusterLess: Deadline-Aware Serverless Workflow Orchestration on Federated Edge Clusters 1312
Reza Farahani, Mario Colosi, Ilir Murturi, Stefan Nastic, Massimo Villari, Schahram Dustdar Radu Prodan

Network Digital Untwinning: Towards Backward Optimization of Digital Twins 1323
Zifan Zhang, Dianwei Chen, Anjun Gao, Manhua Wang, Mingzhe Chen, Minghong Fang, Xianfeng Yang Yuchen Liu

Chimera: An Efficient Multimodal Embodied Inference Framework with Complexity-Aware Edge-Cloud Routing 1334
Siqi Luo, Muen Xue, Liekang Zeng, Tao Ouyang, Pu Wang, Shaoyong Guo, Xu Chen

Panorama: LLM-Guided Heuristics for Digital Twin-Empowered Vehicular Edge Computing 1345
Ziyao Huang, Kui Wu, Weiwei Wu, Xiangtong Qi, Jianping Wang, Jen-Ming Wu

Coordinating Traffic at Group Granularity for Tail Latency Control in Service Meshes 1356
Hokun Park, HyungJun Kim, Hwamin Lee, Heonchang Yu

Task Scheduling and Transmission for Distributed Job Execution 1367
Shangyi Shao, Xueyan Tang

EdgeTP: Enabling Distributed Full-size Large Language Model Inference on Edge Devices with Tensor Parallelism 1378
Laizhong Cui, Yi Guo, Weixuan Peng, Yaodong Huang

PipeNN: A Predictor-Free Pipeline for Energy-Latency Co-Optimization of Heterogeneous Mobile DAG-DNN Inference 1389
Yukun Tian, Tianwei Jiang, Kun Tian, Ruiting Zhou, Yunxin Liu, Zhi Zhou, Fang Dong, Mengyang Liu Wentao Liu

PR-Reduce: Heterogeneity-Aware Proportional Allreduce for Distributed Training 1399
Yi-Ching Kuo, Kate Ching-Ju Lin

From Ranking to Reasoning: Explainable Web API Recommendation via Semantic Reasoning 1410
Zishuo Xu, Dezhong Yao, Yao Wan

Fast-DFL: Flexible Client Switching for Bandwidth-Constrained Hierarchical Decentralized Federated Learning 1421
Zhongkun Wang, Hailiang Yang, Laizhong Cui

GALUPA: Gossiping Aggregated Locking and Output Proofs for Accountability in BFT 1432
Baolin Li, Huiping Sun, Zhong Chen

A Hybrid Blockchain Design Integrating Proof-of-Work and Proof-of-Quantum-Work Consensus 1443
Jiarui Zhang, Yiming Zeng, Yuanyuan Yang

TrustMix: How to Mix Messages in a Mobile Ad-Hoc Network 1454
Yu Shen, Stefanie Roos, Aiswarya Walter

Cappuccino: Cost-Efficient Heterogeneous Lora Fine-Tuning Via Replica-Level Orchestration 1465
Lingxuan Weng, Fei Xu, Xudong Wu, Xinyi Zhang, Shiqi Cheng, Li Chen, Zhi Zhou

Scale: Sensitivity-Aware Federated Unlearning with Information Freshness Optimization for Mobile Edge Computing 1476
Zihao Ding, Beining Wu, Jun Huang