

ATTENDEES

Conference Program

Preliminary schedule for IEEE CCGrid 2026, 18-21 May 2026, Sydney, Australia. This page will be updated as the detailed technical program, room assignments, and speaker information are finalized.

Monday, 18 May 2026

Tutorials and workshops day

| Time | Conf Room 1 | Conf Room 2 | Conf Room 3 | Conf Room 4 |
|---------------|--|----------------------|------------------------|------------------------------|
| 08:00 - 17:00 | Registration: Foyer | | | |
| 09:00 - 10:30 | Tutorial 1 | CIC2 | FL-RCE | Continuum-RI |
| 10:30 - 11:00 | Morning Tea | | | |
| 11:00 - 12:30 | Tutorial 1 | CIC2 | FL-RCE | Continuum-RI |
| 12:30 - 14:00 | Lunch | | | |
| 14:00 - 15:30 | Tutorial 2 | CIC2 | HEAD | Continuum-RI |
| 15:30 - 16:00 | Afternoon Tea | | | |
| 16:00 - 17:30 | Research Mentorship | | | Continuum-RI |
| 18:00 - 20:00 | Welcome Reception: Foyer | | | |

Session Details

Tutorial and workshop details below were provided via the published tutorial and workshop programmes.

09:00 -
10:30

Tutorial 1: Harnessing HPC and AI for Interactive Analysis and Visualization of Large Scientific Datasets with the National Science Data Fabric (NSDF)

[Tutorial overview and materials](#)

Instructors: Prof. Michela Taufer and Prof. Valerio Pascucci

Session 1 of 2, 09:00 - 10:30

IEEE Workshop on Computational Intelligence in Cluster, Cloud, and Internet Computing (CIC²)

PredictScale: Proactive Autoscaling via Workload Forecasting for Cloud-Edge Microservices

Bablu Kumar, Anshul Verma, Pradeepika Verma and Rajkumar Buyya

09:00 - 09:20

Comparative Analysis of Model Pruning Approaches for Optimizing Small Language Models

Sneha Benny, Christina Terese Joseph and Manu Madhavan

09:20 - 09:40

Mango Leaf Health Classification for Drone Images Using Deep Learning in Cloud

Prince Ganeshwala, Dhruv Gohil, Sumedha Arora, Jaiprakash Verma and Bishwajit Poddar

09:40 - 10:00

Towards Monitoring-Driven and Workload-Aware Scheduling for Federated Learning in Cloud-Edge Environments

Pierluigi Dell'Acqua, Eleonora Di Pietro, Francesco La Rosa, Angelo Marchese, Orazio Tomarchio and Massimo Villari

10:00 - 10:20

Session 1 Wrap-Up / Transition

10:20 - 10:30

Workshop on Federated Learning for Resource Constraint Environments (FL-RCE)

Welcome

09:00 - 09:15

Edge-Native Federated LiDAR Analytics for Robust Guardrail Monitoring in Rural Environments

Maria Teresa Reggio, Francesco La Rosa, Pierluigi Dell'Acqua and Massimo Villari

[Paper Presentation 1](#), 09:20 - 09:40

SpotFed: Cost-Efficient Federated Learning with Spot Instance Compensation

Myungjun Son, Woohyeon Baek, Mehrdad Mahdavi and Mahmut Kandemir

[Paper Presentation 2](#), 09:45 - 10:05

Enhancing Fairness in Federated Learning Using Client Clustering and Personalized Models

Asifa Shah, Kenan Matawie and Bahman Javadi

[Paper Presentation 3](#), 10:10 - 10:30

Research Infrastructures for Experimenting across the HPC-Cloud-Edge Continuum (Continuum-RI)

Welcome

09:30 - 09:40

Everything, Everywhere, All the Time: End-to-end Software Infrastructure for Real-time Coupling of HPC, AI, and IoT

Rich Wolski

[Keynote](#), 09:40 - 10:30

11:00 -
12:30

Tutorial 1: Harnessing HPC and AI for Interactive Analysis and Visualization of Large Scientific Datasets with the National Science Data Fabric (NSDF)

[Tutorial overview and materials](#)

Instructors: Prof. Michela Taufer and Prof. Valerio Pascucci

Session 2 of 2, 11:00 - 12:30

IEEE Workshop on Computational Intelligence in Cluster, Cloud, and Internet Computing (CIC²)

Tutorial

Tutorial Speakers: Dr. Navid Yazdanjue and Prof. Amir H. Gandomi

11:00 - 12:00

A Federated Edge-AI Smart City Framework for Early Parkinson's Disease Detection Using Handwriting and Multimodal Digital Biomarkers

Mahtab Shahin, Riina Palu, Ralf-Martin Soe, Mohammad H. Nadimi-Shahraki and Amir H. Gandomi

12:00 - 12:25

Session 2 Wrap-Up

12:25 - 12:30

Workshop on Federated Learning for Resource Constraint Environments (FL-RCE)

Panel Discussion

11:00 - 12:30

Research Infrastructures for Experimenting across the HPC-Cloud-Edge Continuum (Continuum-RI) Cloud-Native Distributed Data Layer for the Computing Continuum with Integrated Location-Aware Workflows

Michal Orzechowski, Łukasz Opióła, Bartosz Kryza, Bartosz Baliś, Lukasz Dutka, Renata G. Słota and Jacek Kitowski

11:00 - 11:30

Energy Malleable Computing Continuum

Carlos Jaime Barrios Hernandez, Frédéric Le Mouël and Oscar Carrillo

11:30 - 12:00

Towards IoT Rejuvenation: From Machine to Federated Learning Based Regression to Detect Sensor Ageing in Cloud-Edge Continuum

Antonio Celesti, Giovanni Lonia, Maria Fazio, Lyubomir Gotsev, Eugenia Kovatcheva, Denis Chikurtev, Antonino Quattrocchi, Roberto Montanini and Massimo Villari

12:00 - 12:30

14:00 -
15:30

Tutorial 2: Federated Learning for Edge Computing

[Tutorial overview and materials](#)

Instructors: Dr Nan Yang, A/Prof Rodrigo N. Calheiros, Prof Massimo Villari, Mark Gambito, and Prof Bahman Javadi

14:00 - 15:30

IEEE Workshop on Computational Intelligence in Cluster, Cloud, and Internet Computing (CIC²)

InfiTRON: An LLM-Powered Infrastructure for Monitoring, Intervention, and Technology Repair Operations Network

Timothy Tron and Abdelbaset Hamza

14:00 - 14:20

Unified Metadata Architecture and Workflow Automation for SEM Imaging in Support of FAIR and Open Science

Phyo Thandar Thant, Takashi Endo, Yuko Mori, Naomi Hirai, Masashi Takei, Takuya Yumine, Masaharu Munetomo, and Yasutaka Matsuo

[14:20 - 14:40](#)

About the Influence of Workflow Topology on Task Intensity Prediction through Graph Learning

Max Simon Otto, Haci Ismail Aslan, Joel Witzke, Jonathan Bader and Odej Kao

[14:40 - 15:00](#)

A Hybrid Multi-task Network Traffic Classification Model Based on Transformer and Mamba

Baiyang Zhou, Aniket Mahanti, Sudheer Kumar Battula and Ranesh Naha

[15:00 - 15:20](#)

Session 3 Wrap-Up / Closing Remarks

[15:20 - 15:30](#)

Healthcare Evolution through AI and Digital Twins (HEAD)

Session Welcome

[14:00 - 14:05](#)

Reproducible Analytics Workflow for Longitudinal Healthcare Digital Twins

Kunal Rehani and Kenan Matawie

[14:05 - 14:25](#)

Acoustic Features as Potential Vocal Biomarkers for Multiple Sclerosis

Salvatore Giugliano, Ilaria Basile, Raffaele Dubbioso, Rosa Iodice, Lucia Aruta and Giovanna Sannino

[14:25 - 14:45](#)

A Systems-Level Evaluation of Scalable Multimodal Architectures for Social Media Suicide Risk Detection

Vijayalakshmi Saravanan, Sai Karthik Navaluru, Khaled Z. Ibrahim and Lakshman Tamil

[14:45 - 15:05](#)

Physics-Informed Prediction of Uncertain Glucose Dynamics: Integrating Wong-Zakai Corrected Stochastic Differential Equations with Machine Learning

Rossella Laudani, Armando Ruggeri and Massimo Villari

[15:05 - 15:25](#)

Closing and Best Paper Award

[15:25 - 15:30](#)

Research Infrastructures for Experimenting across the HPC-Cloud-Edge Continuum (Continuum-RI)

Research Data Storage in the RI Continuum

David Abramson

[Invited Talk, 14:00 - 15:00](#)

Continuum-Aware Storage: Late-Bound Convergence for Unified Edge-to-Cloud Research Data Infrastructures

Jake Carroll, David Abramson and Bronis R. de Supinski

[15:00 - 15:30](#)

16:00 -
17:30

**Research Infrastructures for Experimenting across the HPC-Cloud-Edge Continuum (Continuum-RI)
Failure Risks and Mitigations for a Wide-Area Digital Nervous System**

Nicolas Erdody

[Invited Talk](#), 16:00 - 17:00

Abstractions, Architecture and Implementation of a Modular IoT Edge Testbed

Loïc Guégan, Salma Tofaily and Issam Raïs

17:00 - 17:30

Tuesday, 19 May 2026

Main conference opening and technical sessions

| Time | Conf Room 1 | Conf Room 2 | Conf Room 3 | Conf Room 4 |
|---------------|---|---------------|-----------------|---------------------------|
| 08:00 - 17:00 | Registration: Foyer | | | |
| 09:00 - 09:30 | Opening Ceremony | | | |
| 09:30 - 10:30 | Keynote: Professor Song Guo Exploring Edge Physical Intelligence: Extreme Quantization, Limitless Memory, and Rapid Evolution | | | |
| 10:30 - 11:00 | Morning Tea | | | |
| 11:00 - 12:30 | GPU-1 | LLM-1 | WORKFLOW | Doctoral Symposium |
| 12:30 - 14:00 | Lunch | | | |
| 14:00 - 15:30 | GPU-2 | LLM-2 | AI-SEC | Posters |
| 15:30 - 16:00 | Afternoon Tea | | | |
| 16:00 - 17:30 | HPC-SCH | AI-SYS | GREEN | |
| 17:30 - 18:30 | Panel Discussion | | | |

Session Details

11:00 - 12:30

GPU Communication and HPC Acceleration (GPU-1)

Session Chair: Pierre Jacquet (ETS Montreal)

Kernel-Initiated One-Sided Networking for GPU-Accelerated AI Workloads

Khaled Hamidouche, John Bachan, Pak Markthub, Peter-Jan Gootzen, Elena Agostini, Sylvain Jeaugey, Aamir Shafi, Georgios Theodorakis and Manjunath Gorentla Venkata

[Full paper](#)

Mammoth: Macro-Level MPI Offloading to Off-path Accelerator in DPU

Jong-Bin Lee, Pu-Rum Seo, Ki-Moon Jeong and Hyun-Wook Jin

[Full paper](#)

Scalable, Topology- and Multi-HCA-Aware Hierarchical GPU Allgather Using Parallel Rings

Amirreza Barati Sedeh, Ryan Grant and Ahmad Afsahi

[Full paper](#)

Efficient LLM Inference Systems - 1 (LLM-1)

Session Chair: Mohsen Amini Salehi (University of North Texas)

Characterizing LLM Inference Energy-Performance Tradeoffs across Workloads and GPU Scaling

Paul Joe Maliakel, Shashikant Ilager and Ivona Brandic

[Full paper](#)

Best Student Paper Nominee

Quicktopia: Iteration-Level GPU Frequency Control for Energy-Latency Co-Optimization in LLM Inference

Soyang Baek, Bodon Jeong, Hongsu Byun and Sungyong Park

[Full paper](#)

ARKV: Adaptive Resource-Efficient KV Cache Management for Long Context LLM Inference under Memory Constraints

Jianlong Lei and Shashikant Ilager

[Full paper](#)

Distributed Workflow Management and Optimization (WORKFLOW)

Session Chair: Rodrigo Calheiros (Western Sydney University)

ElastiFlow: Elastic Resource Management for Iterative Scientific Workflows in Hybrid HPC-Cloud Infrastructure

Srishti Dasgupta, Kavitha Subramaniam and Michael Gerndt

[Full paper](#)

Efficiently Reproducing Distributed Workflows in Notebook-based Systems

Talha Azaz, Raza Ahmad, Saiful Islam, Douglas Thain and Tanu Malik

[Full paper](#)

Compass: Optimizing Compound AI Workflows for Dynamic Adaptation

Milos Gravara, Juan Luis Herrera and Stefan Nastic

[Full paper](#)

Doctoral Symposium

Session Chair: Mohammad Goudarzi (Monash University)

Topology-Aware Task and Data Placement for Distributed Scientific Workflows

Sami Kharma, Tobias Wies, Björn Scheuermann and Florian Schintke

[Doctoral Symposium paper](#)

Energy-Efficient Federated Learning Approaches for Orbital Edge Computing

Bara'Ah Al-Blewi, Bahman Javadi and Rodrigo N. Calheiros

[Doctoral Symposium paper](#)

Accurate Synthetic Tasks for Scientific Workflow Benchmarks

Tobias Wies, Sami Kharma, Tobias Meuser and Björn Scheuermann

[Doctoral Symposium paper](#)

**14:00 -
15:30**

GPU Resource Management and Runtime Optimization (GPU-2)

Session Chair: Nick Brown (The University of Edinburgh)

ADVICE: Automatic Identification of Variables to Checkpoint through Compiler Augmentation

Xin Huang, Luanzheng Guo, Nathan Tallent and Kento Sato

[Full paper](#)

Enhanced SVM for Improving Application Performance under GPU Memory Oversubscription

Bennett Cooper, Thomas Scogland and Rong Ge

[Full paper](#)

Priority-Aware GPU Co-Scheduling for High Performance Computing

Naman Kulshreshtha, Tapasya Patki, Aniruddha Marathe, Tom Scogland and Rong Ge

[Full paper](#)

Efficient LLM Inference Systems - 2 (LLM-2)

Session Chair: Adel Toosi (The University of Melbourne)

EmbdC: Error-Bounded Lossy Video Embedding Compression For On-Device LLM Inference

Bo Jiang, Taolue Yang, Youyuan Liu, Sheng Di and Sian Jin

[Full paper](#)

LLM-Pilot: SLO-Aware and Cost-Efficient LLM Serving on Public Cloud VM Clusters via Offloading

Jinwoo Kim, Kihyun Kim, Hyunsun Chung, Jihoon Yang, James J. Kim, Dong Li and Youngjae Kim

[Full paper](#)

EAFAL: An Edge-Based Agentic Framework for Adaptive Selection between SLMs and LLMs

Chamara Madarasingha, Prajyot Singh, Redowan Mahmud, Mahbuba Afrin, Aneesh Krishna and Salil Kanhere

[Full paper](#)

Security and Reliability in AI Systems (AI-SEC)

Session Chair: Reza Farivar (University of Illinois)

Exploring Silent Data Corruption as a Reliability Challenge in LLM Training

Anton Altenbernd, Philipp Wiesner and Odej Kao

[Full paper](#) **Best Paper Nominee**

Noise-Aware Misclassification Attack Detection in Collaborative DNN Inference

Shima Yousefi and Saptarshi Debroy

[Full paper](#)

SeCI: A Framework for Self-Certified Identity for Autonomous AI Agents

Nehal Al-Otaiby, Mohammad Hammoudeh and Jameleddine Hassine

[Full paper](#)

Posters

Session Chair: Mohsen Amini Salehi (University of North Texas)

AccTS: Co-Design of OS Scheduling and System Architecture for Accelerator-Rich SoC

Weicong Xie, Qiyu Wu, Xiaoyang Wang, Minjie Fan and Jiwu Shu

[Poster](#)

Learning-Based Command Priority Scheduling for Tail Latency Reduction in Storage Systems

Zilong Liang, Ye Xu, Minjie Fan and Jiwu Shu

[Poster](#)

Network Telescope as a source of actionable information in near real-time

Marek Janiszewski, Jan Adamski, Marcin Rytel, Jakub Koman and Wojciech Sołtys

[Poster](#)

Empirical Analysis of GPU Frequency Behavior Under ML Workloads

Truong-Thanh Le, Hoang-Loc La, Amir Taherkordi, Frank Eliassen, Phuong Hoai Ha and Peiyuan Guan

[Poster](#)

AI-Enhanced CTI for SOCs: Practitioner Priorities from a Multi-Sector Survey

Anna Felkner and Antonio Monte Pegado

[Poster](#)

A High-Performance RAID-on-Chip SoC Design Implemented with 28nm CMOS Technology

Xu Gao, Qiyu Wu, Xiaoyang Wang, Minjie Fan and Jiwu Shu

[Poster](#)

FORE-Twin: Neuro-Symbolic Middleware for Latency-Aware Data Staging in Federated Digital Twin Systems

Jiwoo Han, Yangkoo Lee, Kyoungyun Park and Daesub Yoon

[Poster](#)

Edge-Based Perception-Driven Robotic System for Assistive Patient Care

Kanaka Sai Jagarlamudi, Rhys Tague, Jacob Desmond, Chng Wei Lau, Hao Wu, Anupama Ginige and Oliver Obst

[Poster](#)

16:00 -
17:30

HPC Scheduling and Execution Optimization (HPC-SCH)

Session Chair: Michela Taufer (University of Tennessee, Knoxville)

GASched: Goal-Adaptive Hierarchical Reinforcement Learning for Multi-Objective HPC Job Scheduling

Minsol Choo and Sangyoon Oh

[Full paper](#)

Reducing Backfill Failures from Workload Drift with Lightweight Uncertainty Buffers in HPC Job Scheduling

Jiheon Choi and Sangyoon Oh

[Full paper](#)

Automated Configuration of Power-Management Knobs for Optimal HPC Job Executions

Francesco Antici, Andrea Proia, Ryoma Ohara, Toshihiro Hanawa, Zeynep Kiziltan, Andrea Bartolini and Jens Domke

[Full paper](#)

Best Paper Nominee

LLM / AI Training & Inference Systems (AI-SYS)

Session Chair: Nancy Yang (Western Sydney University)

Reuse-Aware Min-Cost Flow Scheduling for Distributed LLM Training in Hybrid Optical-Electrical Networks

Yifan Yang, Gongming Zhao, Hongli Xu and Huihui Tang

[Full paper](#)

Tula: Optimizing Time, Cost, and Generalization in Distributed Large-Batch Training

Sahil Tyagi and Feiyi Wang

[Full paper](#)

VEX: Scaling HNSW-Based Vector Search with DPU Memory and Parallelism

Kihwan Kim, Hyungsun Yoo, Woojung Kim, Donghyun Min, Myungcheol Lee, Jihoon Yang, Weikuan Yu and Youngjae Kim

[Full paper](#)

Energy, Sustainability & Carbon-Aware Computing (GREEN)

Session Chair: Massimo Villari (University of Messina)

Green or Fast? Learning to Balance Cold Starts and Idle Carbon in Serverless Computing

Bowen Sun, Christos Antonopoulos, Evgenia Smirni, Bin Ren, Nikolaos Bellas and Spyros Lalis

[Full paper](#)

STEAM: Realistic Modeling and Systematic Exploration of Composable Techniques for Sustainable Datacenter

Dante Niewenhuis, Sacheendra Talluri, Alexandru Iosup and Tiziano De Matteis

[Full paper](#)

Spanergy: Energy-aware Distributed Tracing for Microservices

Cesar Perdigao Batista, Denis Conan and Sophie Chabridon

[Full paper](#)

Wednesday, 20 May 2026

Keynote, best-paper sessions, and community events

| Time | Conf Room 1 | Conf Room 2 | Conf Room 3 | Conf Room 4 |
|---------------|--|-------------|-------------------|------------------------|
| 08:00 - 17:00 | Registration: Foyer | | | |
| 09:30 - 10:30 | Keynote: Professor Kate Keahey "This is Not a Testbed": How to Build and Operate Experimental Infrastructure | | | |
| 10:30 - 11:00 | Morning Tea | | | |
| 11:00 - 12:30 | CLOUD | FL-1 | OPT-SYS | SCALE Challenge |
| 12:30 - 14:00 | Lunch (Steering Committee) | | | |
| 14:00 - 15:30 | EDGE-AI | FL-2 | DATA-SYS-1 | ICFEC 1 |
| 15:30 - 16:00 | Afternoon Tea | | | |
| 16:00 - 17:30 | EDGE-IoT | FL-3 | DATA-SYS-2 | ICFEC 2 |
| 19:30 - 22:30 | Conference Gala Dinner | | | |

Session Details

11:00 -
12:30

Cloud Infrastructure and Orchestration (CLOUD)

Session Chair: Luiz Bittencourt (University of Campinas)

Larger Cloud Servers, Fewer Hosts? On the Evolution of VM Sizes in IaaS Platforms

Pierre Jacquet, Camille Coti and Marcos Dias De Assuncao

[Full paper](#)

TelePod: Live Migration for Stateful Containers

Mingjie Yan, Atharva Ranade, Xin Zhang and Kartik Gopalan

[Full paper](#)

Best Student Paper Nominee

Cost-Justified Multi-type Resource Fair Scheduling for Kubernetes

Bo Yan, Sujoy Sikdar and Madhusudhan Govindaraju

[Short paper](#)

XCAgent: Automating Multi-Cloud Deployment of Agentic Workflows on FaaS Platforms

Varad Kulkarni, Vaibhav Jha, Nikhil Reddy, Anand Eswaran, Praveen Jayachandran and Yogesh Simmhan

[Short paper](#)

Federated / Decentralized Learning - 1 (FL-1)

Session Chair: Sangyoon Oh (Ajou University)

A Multi-Armed Bandit-Based Participant Selection Method for Federated Recommendation Systems

Jintao Liu, Mohammad Goudarzi and Adel N. Toosi

[Full paper](#)

FedMO: Mobility-Aware Client Selection in Federated Learning for Drone Delivery Systems

Jiang Yuan, Xiao Liu, Jia Xu, Aiting Yao, Frank Jiang and Xuejun Li

[Full paper](#)

Heterogeneous-Resource-Aware Federated Learning with Intelligent LoRA Allocation and Aggregation

Youye Xie, Yao Lian, Kevin Chen, Abdul Latif, Lingzhi Zhao and Reza Farivar

[Full paper](#)

Scheduling & Systems Optimization (OPT-SYS)

Session Chair: Carlos J. Barrios H. (INRIA)

AdaSched: A Performance-Driven Cluster Scheduler for Deep Learning Workloads using Deep Reinforcement Learning

Han Yin, Jialun Li, Xuan Mo and Weigang Wu

[Full paper](#)

Congestion-Aware Pricing for Fast and Efficient Edge-Cloud Computing

Polyzois Soumplis and Emmanouel Varvarigos

[Full paper](#)

SD-MoE: Scenario-Driven MoE Forecasting for Intelligent Elastic Scaling in Cloud Clusters

Xianzhao Guo, Weipeng Cao, Minxian Xu, Dachuan Li, Chuanfei Xu, Xi Tao and Zhong Ming

[Full paper](#)

SCALE Challenge

Session Chair: Michela Taufer (University of Tennessee, Knoxville)

Logical Scalability in Big Data Architectures: Impact of Data Modeling on the Database Layer

Rosario Napoli, Mark Adrian Gambito, Antonio Celesti, Massimo Villari and Maria Fazio

[SCALE Challenge paper](#)

A Scalable Computing Continuum Framework for Ambient Assisted Living

Jahedul Anowar, Koushikur Islam, Kanaka Sai Jagarlamudi, Bahman Javadi and Rodrigo Calheiros

[SCALE Challenge paper](#)

Breaking All-Reduce Bottlenecks with Asynchronous Influence Maximization at Scale

Shubhendra Pal Singhal, Ajay Mandadi, Akihiro Hayashi, Thomas M Conte and Vivek Sarkar

[SCALE Challenge paper](#)

MemScale: Real-Time Change Detection with Near-Linear Scale-Out on HPC Systems

Vijayalakshmi Saravanan, Sai Karthik Navaluru, Jon Ryan Cole, Khaled Z Ibrahim and Lakshman Tamil

[SCALE Challenge paper](#)

Scaling Real-Time Traffic Analytics on Edge-Cloud Fabrics for City-Scale Camera Networks

Akash Sharma, Pranjal Naman, Roopkatha Banerjee, Priyanshu Pansari, Sankalp Gawali, Mayank Arya, Sharath Chandra Madanu, Arun Josephraj, Rakshit Ramesh, Punit Rathore, Anirban Chakraborty, Raghu Krishnapuram, Vijay Kovvali and Yogesh Simmhan

[SCALE Challenge paper](#)

14:00 -
15:30

Edge Systems for AI Agents (EDGE-AI)

Session Chair: Yogesh Sharma (University of Regina)

Orchestrating WASM-based MCP Tool Runtimes for AI Agents across Edge-Cloud Continuum

Moohyun Song, Hayoung Kim, Kyoohyun Lee, Jae Gi Son and Kyungyong Lee

[Full paper](#)

SwiftBot: A Decentralized Platform for LLM-Powered Federated Robotic Task Execution

Hailu Xu, Simon Zhang, Zhengxiong Li, Shuai Xu, Xiaokun Yang and Fangtian Zhong

[Full paper](#)

DriveCache: On-Board Compute Caching for Scalable Vehicular Edge Computing Networks

Suvarthi Sarkar, Aditya Gupta, Salil Kashyap and Aryabartta Sahu

[Full paper](#)

Federated / Decentralized Learning - 2 (FL-2)

Session Chair: Mohammad Goudarzi (Monash University)

Dual-Distilled Heterogeneous Federated Learning with Adaptive Margins for Trainable Global Prototypes

Fatema Siddika, Md Anwar Hossen, Wensheng Zhang, Anuj Sharma, Juan Pablo Munoz and Ali Jannesari

[Full paper](#)

RDA-CAFL: Reputation-Dynamic and Distillation-based Asynchronous Conflict-Aware Federated Learning for UAV Networks

Jie Li, Pengyang Li, Liming Sun, Xingwei Wang and Yihang Zhang

[Full paper](#)

Evaluating Federated Learning Beyond Simulation: A Deployment-Aware Methodology

Mathis Valli, Cedric Tedeschi, Alexandru Costan, Loic Cudennec and Gabriel Antoniu

[Full paper](#)

Hardware-Aware Storage & Memory Systems (DATA-SYS-1)

Session Chair: Jongmoo Choi (Dankook University)

Making Variable-Size I/O Practical in ZNS SSDs

Sijie Lan, Abutalib Aghayev, Mahmut Kandemir and Umesh Maheshwari

[Full paper](#)

NIO-Cache: Device-Affinitive Page Cache Placement Mechanism for NUMA Systems

Jiazheng Zhang, Xiaoyang Wang and Jiwu Shu

[Full paper](#)

S-MSHR: A Scalable MSHR Architecture Using Cache Tag Data-Ready Bits and Index Queues

Shuang Wu, Xu Zhang, Yibin Xu, Yangyang Zhao, Tianyue Lu and Mingyu Chen

[Full paper](#)

Monitoring and Management in the Computing Continuum (ICFEC 1)

Session Chair: Yu Xiao (Aalto University)

K-Sense: A Non-Invasive eBPF Framework for QoS Inference

Abdullah Muslim, Ali Beiti Aydenlou and Stephan Recker

Cost-Effective Processing of IoT Data in the Computing Continuum

Vasileios Karagiannis, Drazen Ignjatovic, Antonios Iosifidis and Stefan Schulte

LLMEdger: Phase-Aware Model Parallelism Scheduler for LLM Inference on Edge

Xinyang Shen and Lena Mashayekhy

16:00 -
17:30

IoT & Edge Systems Performance (EDGE-IoT)

Session Chair: Yogesh Simmhan (Indian Institute of Science)

Benchmarking Message Brokers for IoT Edge Computing: A Comprehensive Performance Study

Tapajit Chandra Paul, Pawissanutt Lertpongrijikorn, Hai Nguyen and Mohsen Amini Salehi

[Full paper](#) Best Paper Nominee

Volatility-Aware Adaptive Context Caching for Real-Time Context-Aware IoT Applications

Ashish Manchanda, Prem Prakash Jayaraman, Abhik Banerjee and Arkady Zaslavsky

[Short paper](#)

Hierarchical Semi-Supervised Federated Learning for UAV-Enabled Fire Monitoring

Mark Adrian Gambito, Bahman Javadi, Lorenzo Carnevale and Massimo Villari

[Short paper](#)

Synthetic Data Generation for Storage Failure Prediction in Large-Scale Systems

Chandranil Chakrabortii, Ana Solorzano and Devesh Tiwari

[Short paper](#)

Federated / Decentralized Learning - 3 (FL-3)

Session Chair: Shaleeza Sohail (The University of Newcastle)

FedOort: A Fair and Efficient Optimization Method for Federated Learning

Zhaohua Zheng, Junhui Du, Qiquan Chen, Xu Han, Qijun Huang and Jin Zhang

[Full paper](#)

Evidential Trust-Aware Model Personalization in Decentralized Federated Learning for Wearable IoT

Murtaza Rangwala, Richard Sinnott and Rajkumar Buyya

[Full paper](#)

CroSatFL: Energy-Efficient Federated Learning with Cross-Aggregation for Satellite Edge Computing

Nan Yang, Bahman Javadi, Rodrigo Neves Calheiros, David Boland and Philip Leong

[Full paper](#)

Data Systems & Storage Engines (DATA-SYS-2)

Session Chair: Laurent Lefevre (INRIA)

ARC: Adaptive Resource Coordination for Write-Stall Mitigation in LSM-Tree

Guangxun Zhao, Yongjie Zhu, Suhwan Shin, Seehwan Yoo and Jongmoo Choi

[Full paper](#)

Altocumulus: Enabling Efficient Erasure Coding in IPFS

Mohammad Rizk, Shadi Ibrahim and Thomas Lambert

[Full paper](#)

BucketLSM: What, After All, Is the I/O Bottleneck Behind L0 Key-Range Overlap in LSM-Based KV Stores?

Jaewan Park, Kyungwook Min, Sungjin Byeon, Taewan Noh, Hyungi Park, Xubin He, Hongyeon Kim and Youngjae Kim

[Full paper](#)

Best Student Paper Nominee

Efficient On-Device Learning and Model Compression (ICFEC 2)

Session Chair: Stefan Schulte (TU Berlin)

Layer-Wise Weight Sharing for Efficient Transformer on SoCs

Saeed Khalilian Gourtani, Hang Xu, Nirvana Meratnia and Anuj Pathania

CORAL: Covariance-Guided Resource Adaptive Learning for Efficient Edge Inference

Ahmad Nabhaan, Zaki Sukma, Rakandhiya Rachmanto, Muhammad Santrijaji, Byungjin Cho, Arief Setyanto and In Kee Kim

Order-Aware Compression for RF-DETR on Edge Devices: Overcoming Graph Fragmentation and Quantization Instability

Farhan Mahmood, Michalis Karamousadakis, Antonis Porichis, Vishwanathan Mohan and Panagiotis Chatzakos

Thursday, 21 May 2026

Final technical sessions and closing

| Time | Conf Room 1 | Conf Room 2 | Conf Room 3 | Conf Room 4 |
|---------------|---|-----------------|-------------------|----------------|
| 08:00 - 17:00 | Registration: Foyer | | | |
| 09:30 - 10:30 | Keynote: Professor Muhammad Usman Cloud computing and security in the era of quantum processors | | | |
| 10:30 - 11:00 | Morning Tea | | | |
| 11:00 - 12:30 | EDGE-CLOUD | ML-SYS-1 | EMERGE-SYS | ICFEC 3 |
| 12:30 - 14:00 | Lunch | | | |
| 14:00 - 15:30 | SERVERLESS-SEC | ML-SYS-2 | | ICFEC 4 |
| 15:30 - 16:00 | Afternoon Tea | | | |

Session Details

11:00 -
12:30

Edge-Cloud Resource Management for AI Workloads (EDGE-CLOUD)

Session Chair: Anshul Verma (Banaras Hindu University)

Multi-Objective Load Balancing for Heterogeneous Edge-Based Object Detection Systems

Daghash Alqahtani, Maria Rodriguez, Muhammad Aamir Cheema and Adel N. Toosi

[Full paper](#)

Deep Reinforcement Learning-driven Edge Offloading for Latency-constrained XR pipelines

Sourya Saha and Saptarshi Debroy

[Full paper](#)

Decentralized Resource Sharing in Edge-Cloud Federations via Multi-Agent Hierarchical Reinforcement Learning

Panagiotis Kokkinakis, Polyzois Soumplis and Emmanouel Varvarigos

[Full paper](#)

Learning-Based Performance Modeling (ML-SYS-1)

Session Chair: Stefan Schulte (TU Berlin)

CloudFormer: An Attention-based Performance Prediction for Public Clouds with Unknown Workload

Amirhossein Shahbazinia, Darong Huang, Luis Costero and David Atienza

[Full paper](#)

PM2Lat: Highly Accurate and Generalized Prediction of DNN Execution Latency on GPUs

Truong-Thanh Le, Hoang-Loc La, Amir Taherkordi, Frank Eliassen, Phuong Hoai Ha and Peiyuan Guan

[Full paper](#)

Towards Proactive AIOps: Transfer Learning for Unsupervised Anomaly Detection via Bi-LSTM

Autoencoder in the Computing Continuum

Danny De Novi, Lorenzo Carnevale and Massimo Villari

[Full paper](#)

Advances in Emerging Systems and Security (EMERGE-SYS)

Session Chair: Weisheng Si (Western Sydney University)

Efficient Concurrent GHZ State Distribution in Quantum Networks

Charles Cao, Weisheng Si, Jong Choi and Sajal Das

[Full paper](#)

QRAP: Adaptive Quantum-Safe Risk-Aware Prioritization for Cloud Applications

Surabhi Garg, Delton Myalil, Nidhi Singh, Shiv Shankar, Meena Singh Dilip Thakur and Rajan M A

[Short paper](#)

BCRLSecureLink: A Blockchain, Cryptography, and Reinforcement Learning-Based Defense Against Link Discovery Attacks in SDVN

Patikiri Arachchige Don Shehan Nilmantha Wijesekara, Harsha Sandaruwan Gardiyawasam Pussewalage, Kalupahana Liyanage Kushan Sudheera and Geeth Priyankara Wijesiri

[Short paper](#)

Lifting to tensors when compiling scientific computing workloads for AI Engines

Nick Brown and Gabriel Rodriguez-Canal

[Short paper](#)

Intelligent Infrastructure and Multi-Tier Orchestration (ICFEC 3)

Session Chair: Kanaka Sai Jagarlamudi (Western Sydney University)

Multi-Provider Caching in Multi-Tier Fog Networks

Ferdous Sharifi, Young Choon Lee and Shaahin Hessabi

NL-CPS: Reinforcement Learning-Based Kubernetes Control Plane Placement in Multi-Region Clusters

Sajid Alam, Amjad Ullah and Ze Wang

UAV-Enabled Integrated Sensing, Semantic Communication, and Computation: Disaster-Oriented Edge Computing and Sensing

Yaxi Liu, Wencan Mao, Xulong Li, Yu Xiao, Wei Huangfu and Keping Long

14:00 -
15:30

Security and Analysis of Serverless Systems (SERVERLESS-SEC)

Session Chair: Salil Kanhere (UNSW)

Kumo: A Security-Focused Serverless Cloud Simulator

Wei Shao, Chongzhou Fang, Khaled Khasawneh, Setareh Rafatirad and Houman Homayoun

[Full paper](#)

LogLearners: Identifying Compromised AI Functions in Serverless Systems

Adil Bin Bhutto, Erik Elmroth and Monowar Bhuyan

[Full paper](#)

PoliFlow: Inferring Control-Flow Policies from Serverless Workflows

Pedro Escalera, Vitor A. Cunha, Joao P. Barraca, Diogo Gomes and Rui L. Aguiar

[Full paper](#)

ML Systems & Optimization (ML-SYS-2)

Session Chair: Ranesh Naha (Queensland University of Technology)

ACNNS: A Multi-interest Recommendation Model with Capsule Network

Yan Zhang, Xiaotong Cui, Nan Wang and Yingli Zhong

[Full paper](#)

InverseTune: Inverse Synthetic Fine-Tuning for Reliable Structured Output in Small Language Models

Markus Goetz and Hojjat Baghban

[Short paper](#)

SOtrain: Efficient LLM Fine-Tuning on a Consumer GPU via User-space I/O and Scheduling Optimizations

Zhengguo Liu, Hao Lan and Jiwu Shu

[Short paper](#)

RapidGNN: Communication-Efficient Distributed Training on Large-Scale Graph Neural Networks

Arefin Niam, Tevfik Kosar and M. S. Q. Zulkar Nine

[Short paper](#)

Security and Coordination in Federated Learning (ICFEC 4)

Session Chair: Young Choon Lee (Macquarie University)

A Federated LLM-based Framework for DDoS Defense in Mobile Edge Computing

Shuo Zhang, Kousuke Mori and Toshio Hirotsu

Atlas Synchronization in the Hierarchical Federated Learning Continuum

Antonios Iosifidis, Vasileios Karagiannis and Stefan Schulte