IPDPS 2026 PhD Forum

1. Kyrian Adimora (The University of Kansas), *Machine Learning-Based Multi-Objective Optimization for HPC Workload Scheduling: A GNN-RL Approach*
2. Mattia Balla (Politecnico di Milano), *Memory-Centric Design of Hardware Accelerators for Deep Learning Systems*
3. Nicholas Cassarino (University of North Carolina at Charlotte), *Poor Performance of Unified Memory Management*
4. Neelesh Gupta (University of Southern California), *Predictive Data Movement for LLM-Mediated Agentic AI Systems*
5. Md Arafat Hossain (Iowa State University), *Representation Geometry for Large Language Model Training and Post-Training Decisions*
6. Changxin Li (Case Western Reserve University), *Optimizing Deployment of Irregular Sparse Deep Learning Workloads through ILP-Based Scheduling for Low-Latency Inference*
7. Nazreen Shah (IIIT Delhi), *Towards Robust and Adaptive Federated Learning: Personalized Representations, Continual Adaptation, and Communication-Efficient Methods*
8. Ruimin Shi (KTH Royal Institute of Technology), *Vector-length Agnostic Architecture in Post-Moore High-performance Computing*
9. Irene Simó Muñoz (Cornell University), *Low-Precision SpMV and s-step SGD on Processing-in-Memory*
10. Giuseppe Sorrentino (Politecnico di Milano), *Unleashing Heterogeneous Systems Capabilities to Enhance Compute-Intensive Workloads*
11. Tommaso Spagnolo (Politecnico di Milano), *DIMC-centric accelerator architectures for edge AI and scalable low-end HPC*

12. Minqiu Sun (University of Delaware), *Optimizer-Aware Checkpointing for Efficient Fault Tolerance in LLM Post-Training*
13. Kutay Tasci (Case Western Reserve University), *Scalable Geometric Graph Neural Networks: Parallel and Distributed Optimization Strategies*
14. Maya Taylor (University of Illinois at Urbana-Champaign), *Mapping Unstructured Sparse Computation to Dataflow Architectures*
15. Alessandro Verosimile (Politecnico di Milano), *HW-SW co-design techniques for DT-based Ensemble models on embedded systems*
16. Xincheng Yang (Illinois Institute of Technology)*, Achieving Consistency and Automatic Workload Balancing in Scalable Distributed Deployments*