Walid A. Hanafy

Postdoctoral Research Associate

Manning College of Information and Computer Sciences
University of Massachusetts Amherst
(+1) 413 345-9453
whanafy@cs.umass.edu
WebPage *washraf* *whanafy*

Summary

I am a distributed systems researcher, with extensive experience in developing *efficient* and *intelligent* distributed systems. My current research focuses on building *sustainable* computing and societal infrastructures, as well as designing *resource-efficient* and *failure-resilient* AI systems for resource-constrained edge environments.
My work has been published in top-tier conferences, e.g., *SIGMETRICS, ASPLOS, INFOCOM, e-Energy, HPDC, SoCC, IoTDI, SEC, and IWQoS*.

Education

2019–2025 **MS'23/PhD'25, Computer Science**, *University of Massachusetts Amherst*, MA, USA.
Thesis Title Carbon-aware Resource Management for Cloud Computing Platforms.
Supervisor Prof. Prashant Shenoy, LASS research lab.

2015–2018 **Masters of Computer Engineering**, *Faculty of Engineering*, Helwan University, Egypt.
Thesis Title Container Type Virtualization Management in Cloud Computing.
Supervisors Prof. Sameh Salem, Prof. Amr Mohamed.

2011–2012 **Diploma in Information Technology**, *Major in System Development, Minor in Software Architecture*, Information Technology Institute (ITI), Egypt.

2006–2011 **BS of Computer Engineering (ClassValedictorian)**, Helwan University, Egypt.
Thesis Title Assistive Drive System: Driver Behavior Monitoring using computer vision and ultrasonic sensors.
Supervisor Prof. Sameh Salem

Research Experience

2025–Now **Postdoctoral Research Associate**, *University of Massachusetts Amherst*, MA, USA.

Computational Decarbonization, ***NSF CoDec***.
Papers: HPDC'25, HotCarbon'25, COMPASS'25
Designed and implemented approaches to assess and reduce the carbon emissions of computing and other societal activities.

Faulilure Resilient AI.
Papers: SoCC'25, MilCom'25
Developed failure resilient AI serving systems for resource constrained edge devices.

Foundation Models (FMs) and Large Langage Models (LLMs).
Papers: SEC'25, NeurIPS'25/BERT^{2}S
Developed FM services as well as IoT access control systems based on LLMs.

2020–2025 **Research Assistant**, *University of Massachusetts Amherst*, MA, USA.

Sustainable Computing.
Papers: ASPLOS'23/'24, HotCarbon'23/'24, SIGMETRICS'24, e-Energy'24, SoCC'24
Designed and implemented Carbon and Energy aware system monitoring and management solutions that help reduce the carbon footprint of computing.

Model-Serving on The Edge.
Papers: WEEE'21, TAAS'23, IoTDI'23, SEC'23, MilCom'23, MilCom'24
Designed and implemented Model-Serving platforms that adapt to workload and environment dynamics while providing fairness, resiliency, and latency guarantees.

2021, 2022 **Summer Intern**, *Nokia Bell Labs, Network Systems Research Department*, NJ, USA.

Low-Latency Model Serving, *Papers: IWQoS'23*.
Designed a low latency model serving framework using RDMA and GPUDirect RDMA.

Telemery system, *Patent: US11876691B2, Papers: INFOCOM'24*.
Designed an End-to-End telemetry system for RDMA network traffic.

2018 **Visting Scholar**, *Public University of Navarra, Spain*, Funded by ERASMUS+.
Papers: ICENCO'19
Cleanliness detection of solar panels using computer vision and machine learning.

2015–2019 **Research Assistant**, *Helwan University*, Egypt.
Papers: ICENCO'17, ICCES'17, IEEE Access'19
Developed elasticity control/load balancing algorithms for containers as a service cloud.

Honors and Awards

2025 **Krithi Ramamritham Computer Science Scholarship** for Outstanding Graduate Students.

2024 **Best Student Paper Award** for CarbonScaler in *ACM SIGMETRICS/IFIP Performance 2024*.

2024 **NSF Student Travel Grant Award** to attend *ACM SIGMETRICS/IFIP Performance 2024*.

2021 **Outstanding Innovation Award** for my summer internship from *Nokia Bell Labs*, USA.

2012 **ICT4Change - 1st place** for my ITI graduation project from *ICT Ministry*, Egypt.

2011 **Young Innovators Award** for my BS graduation project from *Nahdet El Mahrosa*, Egypt.

Publications and Patents

Selected Publications

2025 Li Wu, **Walid A. Hanafy**, Abel Souza, Khai Nguyen, Jan Harkes, David Irwin, Mahadev Satyanarayanan, and Prashant Shenoy. CarbonEdge: Leveraging Mesoscale Spatial Carbon-Intensity Variations for Low Carbon Edge Computing. In *Proceedings of the 34th International Symposium on High-Performance Parallel and Distributed Computing (HPDC '25)*, 2025.

2025 Phuthipong Bovornkeeratiroj, **Walid A. Hanafy**, David Irwin, and Prashant Shenoy. GreenThrift: Optimizing Carbon and Cost for Flexible Residential Loads. *ACM Journal on Computing and Sustainable Societies*, March 2025.

2024 **Walid A. Hanafy**, Qianlin Liang, Noman Bashir, Abel Souza, David Irwin, and Prashant Shenoy. Going Green for Less Green: Optimizing the Cost of Reducing Cloud Carbon Emissions. In *Proceedings of the 29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 3*, ASPLOS '24, page 479–496, 2024.

2024 **Walid A. Hanafy**, Qianlin Liang, Noman Bashir, David Irwin, and Prashant Shenoy. CarbonScaler: Leveraging Cloud Workload Elasticity for Optimizing Carbon-Efficiency. *SIGMETRICS Performance Evaluation Review*, volume 52, page 49–50, June 2024. **Best Student Paper Award**.

Other Conference and Journal Articles

2025 Li Wu, **Walid A. Hanafy**, Tarek Abdelzaher, David Irwin, Jesse Milzman, and Prashant Shenoy. FailLite: Failure-Resilient Model Serving for Resource-Constrained Edge Environments. In *Proceedings of the 2025 ACM Symposium on Cloud Computing*, SoCC'25, 2025.

2025 Hetvi Shastri, **Walid A. Hanafy**, Li Wu, David Irwin, Mani Srivastava, and Prashant Shenoy. LLM-Driven Auto Configuration for Transient IoT Device Collaboration. In *Proceedings of the Tenth ACM/IEEE Symposium on Edge Computing (SEC)*, 2025.

2025 Diptyaroop Maji, **Walid A. Hanafy**, Li Wu, David Irwin, Prashant Shenoy, and Ramesh K. Sitaraman. Data Centers Carbon Emissions at Crossroads: An Empirical Study. *ACM SIGEnergy Energy Informatics Review*, volume 5, page 48–55, 2025. Presented in HotCarbon'25.

2025 David Irwin, Prashant Shenoy, Mohammad Hajiesmaili, **Walid A. Hanafy**, Jimi Oke, Ramesh Sitaraman, Yuvraj Agarwal, Geoff Gordon, Zico Kolter, Deepak Rajagopal, Mani Srivastava, Vivienne Sze, Priya Donti, Andrew Chien, John Birge, Ali Hortacsu, and Line Roald. A Vision for Computational Decarbonization of Societal Infrastructure. *IEEE Internet Computing*, pages 1–7, 2025.

2024 **Walid A. Hanafy**, Roozbeh Bostandoost, Noman Bashir, David Irwin, Mohammad Hajiesmaili, and Prashant Shenoy. The War of the Efficiencies: Understanding the Tension between Carbon and Energy Optimization. *ACM SIGEnergy Energy Informatics Review*, volume 4, page 87–93, September 2024. Presented in HotCarbon'23.

2024 Jorge Murillo, **Walid A. Hanafy**, David Irwin, Ramesh Sitaraman, and Prashant Shenoy. CDN-Shifter: Leveraging Spatial Workload Shifting to Decarbonize Content Delivery Networks. In *Proceedings of the 2024 ACM Symposium on Cloud Computing*, SoCC '24, page 505–521, 2024.

2024 Jinyang Li, Yizhuo Chen, Tomoyoshi Kimura, Tianshi Wang, Ruijie Wang, Denizhan Kara, Yigong Hu, Li Wu, **Walid A. Hanafy**, Abel Souza, Prashant Shenoy, Maggie Wigness, Joydeep Bhattacharyya, Jae Kim, Guijun Wang, Greg Kimberly, Josh Eckhardt, Denis Osipychev, and Tarek Abdelzaher. Acies-OS: A Content-Centric Platform for Edge AI Twinning and Orchestration. In *2024 33rd International Conference on Computer Communications and Networks (ICCCN)*, pages 1–9, 2024.

2024 Hyunseok Chang, **Walid A. Hanafy**, Sarit Mukherjee, and Limin Wang. INSERT: In-Network Stateful End-to-End RDMA Telemetry. In *IEEE Conference on Computer Communications*, INFOCOM'24, 2024.

2024 Roozbeh Bostandoost, **Walid A. Hanafy**, Adam Lechowicz, Noman Bashir, Prashant Shenoy, and Mohammad Hajiesmaili. Data-driven algorithm selection for carbon-aware scheduling. *ACM SIGEnergy Energy Informatics Review*, volume 4, page 148–153, April 2024. Presented in HotCarbon'24.

2024 Roozbeh Bostandoost, Adam Lechowicz, **Walid A. Hanafy**, Noman Bashir, Prashant Shenoy, and Mohammad Hajiesmaili. LACS: Learning-Augmented Algorithms for Carbon-Aware Resource Scaling with Uncertain Demand. In *The 15th ACM International Conference on Future and Sustainable Energy Systems (e-Energy '24), June 4—7, 2024, Singapore*, 2024.

2023 **Walid A. Hanafy**, Limin Wang, Hyunseok Chang, Sarit Mukherjee, T.V. Lakshman, and Prashant Shenoy. Understanding the Benefits of Hardware-Accelerated Communication in Model-Serving Applications. In *2023 IEEE/ACM 31st International Symposium on Quality of Service (IWQoS)*, IWQoS'23, pages 1–10, 2023.

2023 Abel Souza, Noman Bashir, Jorge Murillo, **Walid A. Hanafy**, Qianlin Liang, David Irwin, and Prashant Shenoy. Ecovisor: A Virtual Energy System for Carbon-Efficient Applications. In *Proceedings of the 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems*, volume 2 of *ASPLOS 2023*, page 252–265, 2023.

2023 Qianlin Liang, **Walid A. Hanafy**, Noman Bashir, David Irwin, and Prashant Shenoy. Energy Time Fairness: Balancing Fair Allocation of Energy and Time for GPU Workloads. In *Proceedings of the 8th ACM/IEEE Symposium on Edge Computing (SEC)*, 12 2023.

2023 Qianlin Liang, **Walid A. Hanafy**, Noman Bashir, Ahmed Ali-Eldin, David Irwin, and Prashant Shenoy. Dělen: Enabling Flexible and Adaptive Model-serving for Multi-tenant Edge AI. In *Proceedings of IEEE/ACM Eighth International Conference on Internet-of-Things Design and Implementation (IoTDI), San Antonio*, 5 2023.

2023 Qianlin Liang, **Walid A. Hanafy**, Ahmed Ali-Eldin, and Prashant Shenoy. Model-driven Cluster Resource Management for AI Workloads in Edge Clouds. *ACM Transactions on Autonomous and Adaptive Systems*, volume 18, pages 1–26, 2023.

2019 **Walid A. Hanafy**, Alfredo Pina, and Sameh A. Salem. Machine Learning Approach for Photovoltaic Panels Cleanliness Detection. In *2019 15th International Computer Engineering Conference (ICENCO)*, pages 72–77, 2019.

2019 **Walid A. Hanafy**, Amr E. Mohamed, and Sameh A. Salem. A New Infrastructure Elasticity Control Algorithm for Containerized Cloud. *IEEE Access*, volume 7, pages 39731–39741, 2019.

2017 **Walid A. Hanafy**, Amr E. Mohamed, and Sameh A. Salem. Novel Selection Policies for Container-based Cloud Deployment Models. In *2017 13th International Computer Engineering Conference (ICENCO)*, pages 237–242, 2017.

2017 **Walid A. Hanafy**, Amr E. Mohamed, and Sameh A. Salem. A Load Balancing with Power Optimization Algorithm for Container-based Infrastructure Management. In *2017 12th International Conference on Computer Engineering and Systems (ICCES)*, pages 161–166, 2017.

Workshops and Posters

2025 Hetvi Shastri, **Walid A. Hanafy**, Li Wu, David Irwin, Mani Srivastava, and Prashant Shenoy. *Poster Abstract: Rethinking Collaboration Among Mobile Devices in IoT Environments*, page 676–677. 2025.

2025 Hetvi Shastri, Pragya Sharma, **Walid A. Hanafy**, Mani Srivastava, and Prashant Shenoy. FMTK: A Modular Toolkit for Composable Time Series Foundation Model Pipelines. In *The NeurIPS workshop on Recent Advances on Time Series Foundation Models: Have We Reached the BERT Moment? (BERT^{2}S)*, 2025.

2024 Li Wu, **Walid A. Hanafy**, Abel Souza, Tarek Abdelzaher, Gunjan Verma, and Prashant Shenoy. Enhancing Resilience in Distributed ML Inference Pipelines for Edge Computing. In *MILCOM 2024 - 2024 IEEE Military Communications Conference (MILCOM)*, pages 1–6, 2024.

2024 Mihir Shenoy and **Walid A. Hanafy**. Home Decarbonizer: Greening Household Energy Consumption Using Temporal Shifting. In *2024 IEEE MIT Undergraduate Research Technology Conference (URTC)*, pages 1–5, 2024.

2023 **Walid A. Hanafy**, Li Wu, Tarek Abdelzaher, Suhas Diggavi, and Prashant Shenoy. Failure-Resilient ML Inference at the Edge through Graceful Service Degradation. In *Proceedings of the 41st IEEE Military Communications Conference (MILCOM) workshop on Internet of Things for Adversarial Environments*, 10 2023.

2021 **Walid A. Hanafy**, Tergel Molom-Ochir, and Rohan Shenoy. Design Considerations for Energy-Efficient Inference on Edge Devices. In *Proceedings of International Workshop on Energy-Efficient Learning at the Edge*, WEEE'21, page 302–308, 2021.

Patents

2023 Hyunseok Chang, Limin Wang, Sarit Mukherjee, and **Walid Abdelrahman**. End-to-End RDMA Telemetry System, 2023. Patent US11876691B2.

Under Review

2026 Jorge Murillo, **Walid A. Hanafy**, David Irwin, Ramesh K. Sitaraman, and Prashant Shenoy. Go with the Flow: Analyzing the Carbon Footprint of Green Streaming, 2026.

2026 Fabian Francis Mkocheko, **Walid A. Hanafy**, Adam Lechowicz, David Irwin, and Prashant Shenoy. City-Scale Decarbonization of Residential Heating under Transformer Constraints, 2026.

2026 Roozbeh Bostandoost, Adam Lechowicz, **Walid A. Hanafy**, Prashant Shenoy, and Mohammad Hajiesmaili. Quantifying the Carbon Reduction of DAG Workloads: A Job Shop Scheduling Perspective, 2026.

2025 **Walid A. Hanafy**, Li Wu, David Irwin, and Prashant Shenoy. CarbonFlex: Enabling Carbon-aware Provisioning and Scheduling for Cloud Clusters, 2025.

2025 **Walid A. Hanafy**, Thanathorn Sukprasert, Abdel Souza, David Irwin, and Prashant Shenoy. Untangling the Carbon-Cost Tradeoffs and Stampede Effect Challenges in Cloud Computing, 2025.

2025 Krishna Praneet Gudipaty, **Walid A. Hanafy**, Kaan Ozkara, Qianlin Liang, Jesse Milzman, Prashant Shenoy, and Suhas Diggavi. MEL: Multi-level Ensemble Learning for Resource-Constrained Environments, 2025.

Conference Presentations & Talks

July 2025 **HotCarbon'25**, Cambridge, MA, USA.
Data Centers Carbon Emissions at Crossroads: An Empirical Study

March 2025 **Guest Lecturer**, UMass COMPSCI 677 — Distributed and Operating Systems.
A Hitchhiker's Guide to Sustainable Computing.

Feb. 2025 **NESysDay**, Northeastern University, Boston, MA, USA.
Carbon-aware Scheduling: Approaches and Trade-offs.

Dec. 2024 **Guest Lecturer**, UMass COMPSCI CS590X — Decarbonization and Data Science.
Introduction to Carbon-aware Computing.

July 2024 **HotCarbon'24**, Santa Cruz, CA, USA.
Data-driven Algorithm Selection for Carbon-Aware Scheduling. Santa Cruz, CA, USA.

June 2024 **SIGMETRICS'24**, Venice, Italy.
CarbonScaler: Leveraging Cloud Workload Elasticity for Optimizing Carbon-Efficiency.

May 2024 **ASPLOS'24**, San Diego, CA, USA.
Going Green for Less Green: Optimizing the Cost of Reducing Cloud Carbon Emissions.

July 2023 **MILCOM'23**, Boston, MA, USA.
Failure-Resilient ML Inference at the Edge through Graceful Service Degradation.

July 2023 **HotCarbon'23**, Boston, MA, USA.
The War of the Efficiencies: Understanding the Tension between Carbon and Energy Optimization.

June 2023 **IWQoS'23**, Orlando, FL, USA.
Understanding the Benefits of Hardware-Accelerated Communication in Model-Serving.

June 2021 **e-Energy WEEE'21**, Online.
Design Considerations for Energy-Efficient Inference on Edge Devices.

Dec. 2017 **ICENCO 2017**, Cairo, Egypt.
Novel Selection Policies for Container-based Cloud Deployment Models.

Dec. 2017 **ICCES'12**, Cairo, Egypt.
A Load Balancing with Power Optimization Algorithm for Container-based Infrastructure.

Teaching Experience

University of Massachusetts Amherst, Teaching Assistant

Spring 2023 COMPSCI 677: Distributed and Operating Systems

Spring 2020 COMPSCI 577: OS Design and Implementation

Information Technology Institute (ITI), Visiting Lecturer

2015-2019 Design and teach courses in Software Design, Cloud Computing, and Software Architecture

Technologies: Microsoft Stack, Microsoft Azure, and Docker *multiple recurrent courses*

Helwan University, Teaching Assistant

2015-2019 Software Engineering, Database Systems, Operating Systems, and Distributed Systems

Other Work Experience

2017-2019 Testicide (start-up), Egypt, Research Engineer

Led the design and implementation of a "codeless" GUI testing framework.

Tools: Azure Service Fabric, CosmosDB, AngualrJS

2016-2017 IT-Bits, UAE, Software Engineer

Led the successful development of software solutions for multiple pharmaceutical companies.

Tools: asp.net Core, MSSQL server, Angualr 2

2012-2013 Consukorra-PES, Egypt, Software Engineer

Design and Implement a GIS-based data collection platform, leading to a 200% speedup in the collection process.

Tools: QGIS, PostgreSQL, .net framework

Service and Volunteer Work

2021 - Now Journal Reviewer.
IEEE Access, Springer Automated Software Engineering, ACM Computing Surveys, IEEE Internet Compting, IEEE Transactions on Computers, IEEE Transactions on Consumer Electronic

2025 Program Committee Memeber, ACM e-Energy 2026, Banff, Canada.

2025 Program Committee Memeber, IEEE/ACM SEC 2025, Washington, D.C., USA.

2025 Program Committee Memeber, HotCarbon'2025, Cambridge, MA, USA.

2025 Organization Committe Memeber, CarbonMetrics 2025, Stony Brook, NY, USA.

2024-2025 Shadow PC Memeber, Eurosys'24, Eurosys'25.

2024-2025 UMass CICS Turing Program, *Summer Camp for HighSchool Students*, Instructor.

2020-2024 UMass MSA, *Grdauates Branch*, Treasurer.

2021 Artifcat Evaluation Committe Memeber, *Eurosys 2021*.

2017 Docker Egypt Meetup, *Tutorials on containers*, Funded by Microsoft and Docker, Egypt.

Technical skills

Programming Python, C/C++, C#, Java

Web HTML 5, .net Stack, AngularJS

Database MSSQL, PostgreSQL, MySQL, COSMOSDB

Cloud VMware, AWS, Microsoft Azure, Docker, and K8s

Testing XUnit, Specflow, and Selenium

ML Python Stack and Pytorch

Referees

Prashant Shenoy, *University of Massachusetts Amherst* ✉ shenoy@umass.edu
David Irwin, *University of Massachusetts Amherst* ✉ deirwin@umass.edu
Ramesh K. Sitaraman, *University of Massachusetts Amherst* ✉ ramesh@cs.umass.edu
Mohammad Hajiesmaili, *University of Massachusetts Amherst* ✉ hajiesmaili@cs.umass.edu
Ahmed Ali-Eldin Hassan, *Chalmers University of Technology* ✉ ahmed.hassan@chalmers.se

Updated On: September 29, 2025