Peng Kang

6016 J Street, Riverview Hall, Room 5044, Sacramento, CA 95819-2635
peng.kang@csus.edu | https://pengkang12.github.io

Research Interests

Cloud/Edge Computing, Operating System, and Applied AI for System

Working Experience

California State University, Sacramento *08/2024 - Now*
Assistant Professor in the Department of Computer Science

Google, Pittsburgh *05/2022 - 08/2022*
Software Engineer Intern
- Mentors: Tom Black, Max Glick

Jianxun Culture, Shanghai *01/2018 - 07/2018*
Software Development Engineer

Baidu, Beijing *10/2016 - 04/2017*
Site Reliability Engineer

Education

The University of Texas at San Antonio
Ph.D. in Computer Science *2018 - 2024*
- Dissertation: *SLO-Aware Resource Management for Edge Computing*
- Supervisor: Dr. Palden Lama

The University of Texas at San Antonio
M.S. in Computer Science *2018 - 2023*

Xi'an Microelectronic Technology Institute
M.S. in Computer Science *2013 - 2016*
- Thesis: *Research on high reliability embedded real-time operating system*
- Supervisor: Prof. Xubang Shen

Nanjing University of Aeronautics and Astronautics
B.S. in Electrical Engineering *2009 - 2013*

Publications

Conference Publications

• Data-priority Aware Fair Task Scheduling for Stream Processing at the Edge (**Selected as the best paper**).
Faiza Akram, **Peng Kang**, Palden Lama, Samee U. Khan
In *the 8th IEEE Cloud Summit, Washington, DC, USA*, 2024.

• Enhanced Converting Autoencoder based Framework for Low-latency Energy-efficient DNN with SLO-constraints.
Hasanul Mahmud, **Peng Kang**, Kevin Desai, Palden Lama and Sushil Prasad
In *the 8th IEEE Cloud Summit, Washington, DC, USA*, 2024.

- High-throughput Real-time Edge Stream Processing with Topology-Aware Resource Matching.
  Peng Kang, Samee U. Khan, Xiaobo Zhou, and Palden Lama
  In *the 24nd IEEE International Symposium on Cluster, Cloud and Internet Computing (CCGrid)*, 2024.
- A Converting Autoencoder Toward Low-latency and Energy-efficient DNN Inference at the Edge.
  Hasanul Mahmud, **Peng Kang**, Kevin Desai, Palden Lama and Sushil Prasad
  In the *6th Workshop on Parallel AI and Systems for the Edge (PAISE)*, 2024.
- Some New Observations on SLO-aware Edge Stream Processing.
  Amna Shahid, **Peng Kang**, Palden Lama, and Samee U. Khan
  In *IEEE Cloud Summit* 2023.
- Kneescale: Efficient Resource Scaling for Serverless Computing at the Edge.
  Xue Li, **Peng Kang**, Jordan Molone, Wei Wang, and Palden Lama
  In *the 22nd IEEE International Symposium on Cluster, Cloud and Internet Computing (CCGrid)*, 2022.
- SLO-Aware Virtual Rebalancing for Edge Stream Processing.
  Peng Kang, Palden Lama, and Samee U. Khan
  In *the 9th IEEE International Conference on Cloud Engineering (IC2E)*, 2021.
- Robust Resource Scaling of Containerized Microservices with Probabilistic Machine Learning.
  Peng Kang and Palden Lama
  In *the 13th IEEE/ACM International Conference on Utility and Cloud Computing (UCC)*, 2020.

Journal Publications

- Multicore embedded real-time scheduling algorithm based on gang scheduling.
  Peng Kang, Congxiu Liu, and Xubang Shen
  Microelectronics and Computer, 2016.

Under Review and In Preparation

- Adaptive Performance Modeling for Edge Stream Processing System.
  Peng Kang, Faiza Akram, Palden Lama, Samee U. Khan
  Target to: *Journal of Parallel and Distributed Computing*, 2024.

Teaching Experience

Fall 2024	**CSC/CPE 159 Operating System Pragmatics**, Lecturer
	CSC 190 Senior Project, Lecturer
Spring 2024	**CS 4613 Senior Design**, Lecturer
	CS 4843 Cloud Computing, Teaching assitant
Fall 2023	**CS 3423 System Programming Lab Recitation**, Lecturer
Fall 2022 - Spring 2023	**CS 5573 Cloud Computing**, Teaching assitant
Summer 2019	**CS 3843 Computer Organization Lab Recitation**, Lecturer
Fall 2018	**CS 3733 Operating System**, Teaching Assistant

Awards & Honors

2024	**Graduate Student Professional Development Award**, UTSA
2024	**CCGrid 2024 travel grant**, NSF
2022	**Who's Who**, UTSA

2021 **NSDI'21 Student Grant**, The 18th USENIX Symposium on Networked Systems Design and Implementation (NSDI '21)

Alvarez Research Competitive Scholarship, UTSA

2020 **Phi Kappa Phi**, Honor Society

2008 **Provincial 2nd Prize**, National High School Mathematics League (Gansu, China)

Presentations and Talks

High-throughput Real-time Edge Stream Processing with Topology-Aware Resource Matching. CCGrid, 2024.

SLO-Aware Virtual Rebalancing for Edge Stream Processing. IC2E, 2021.

Robust Resource Scaling of Containerized Microservices with Probabilistic Machine Learning. UCC, 2020.

UTSA AI Summit, 2019.

UTSA Computer Science Research Showcase, 2019, 2022.

Professional Services & Activities

Reviewer

2024 IEEE International Conference on Data Mining (ICDM)
2023 IEEE Transactions on Network Science and Engineering (TNSE)
2022 IEEE International Conference on Communications (ICC)

Web master

2020 - 2024 IEEE Computer Society Technical Committee on Distributed Processing (TCDP)

Session Chair

2024 IEEE International Symposium on Cluster, Cloud and Internet Computing

Professional Memberships

2019 - Now IEEE Member
2024 - Now California Faculty Association

Mentoring

- 05/2023 - 06/2024
  Faiza Akram, PhD Student, Mississippi State University
  Project: *Explore data distribution at edge stream processing.*
- 06/2022 - 08/2023
  Amna Shahid, Master Student, Mississippi State University, Graduated
  Project: *Observation of data priority at edge stream processing.*

Certificates

2024 NSF AI Spring School

Google Project Management

Google IT Automation with Python

Technical Skills

Python: Django/Tornado, Postgres/Mysql, Memcached/Redis, Celery/RabbitMQ, RESTful, Machine Learning (Scipy, Scikit-learn, Pandas, Keras, Tensorflow, PyTorch).

Java: Stream processing (Apache Storm, Apache Spark).

C/C++: OpenMP, Embedded OS (VxWorks), Linux system development.

Cloud: Google Cloud, AWS, KVM, Ubuntu/CentOS, Kubernetes, Docker, Microservices.

Miscellany: shell, git, Jenkins, Jetson Nano, Raspberry Pi, Project Management.