ASST. PROF. RACHATA AUSAVARUNGNIRUN (HAVE GREEN CARD)

Thai-German Graduate School, KMUTNB, Bangkok, Thailand
r.ausavarungnirun@gmail.com

Research Interest

Computer Architecture: Heterogeneous multi-core architecture, GPU architecture, GPU in the cloud, Processing-in-Memory, Network-on-chip, Memory subsystem, Quality of Service
Computer System: Distributed system, IoT, Disaggregated memory, Containerization, Virtualization, Storage system

Education

Ph.D. in Electrical and Computer Engineering — Spring 2017
Carnegie Mellon University, Pittsburgh, PA
- Thesis: Techniques for Shared Resource Management in Systems with Throughput Processors

Bachelor of Science in Electrical and Computer Engineering
And Bachelor of Science in Computer Science (Dual Degree)
Carnegie Mellon University, Pittsburgh, PA — May 2010
- GPA: 3.93

Work Experience

The Sirindhorn Thai-German Graduate School (TGGS) — Summer 2019 -- Now
King Mongkut's University of Technology North Bangkok
- Assistant Professor (Tenured)

Carnegie Mellon University, Pittsburgh, PA — Spring 2017 – Summer 2019
Department of Electrical and Computer Engineering
- Post Doctoral Researcher (PI: Onur Mutlu)
  - Research focus: QoS for Heterogeneous Architecture, Virtualizing GPUs, Processing in Memory, NUMA on GPUs, Memory Subsystems

Research Grants

National Science and Technology Development Agency Grant for the Development of High-quality Graduates (009/2565)
- Main PI: High-performance Serverless Platforms for Modern Workloads
- 3,030,000 THB Awarded in 2022

Thailand Ministry of Higher Education, Science, Research and Innovation Research Grant for New Scholar (RGNS 64-091)
- Main PI: Handling Oversubscriptions in Unified Virtual Memory for Multi-tenant GPUs
- 600,000 THB Awarded in 2021

National Science and Technology Development Agency Grant for the Development of High-quality Graduates (013/2563)
- Main PI: A Novel Graph Processing Framework for High-performance Computing
- 787,000 THB Awarded in 2020

Internships

VMWare, Palo Alto, CA — Summer 2015 - Spring 2016
VMWare Research Group
- Research Intern (Mentor: Chris Rossbach)
  - Research focus: QoS and Performance Isolation for GPUs

Intel Corporation, Hillsboro, OR — Summer 2013
Intel Labs
- Research Intern (Mentor: Micheal Mesnier)
  - Research focus: QoS for Virtual Machines

NVIDIA, Santa Clara, CA — Summer 2012
NVIDIA Research
- Research Intern (Mentor: Albert Meixner)
  - Research focus: GPU microarchitectures and memory microarchitectures

Advanced Micro Devices (AMD), Bellevue, WA — Summer 2011
Research and Development Lab (RADL)
- Co-op engineer (Mentor: Gabriel Loh)
  - Research focus: Resource management for a CPU/GPU system

ASST. PROF. RACHATA AUSAVARUNGNIRUN (HAVE GREEN CARD)

Thai-German Graduate School, KMUTNB, Bangkok, Thailand
r.ausavarungnirun@gmail.com

Patent

Rachata Ausavarungnirun, Gabriel Loh, Method and Apparatus for Batching Memory Request. US Patent #8775762 B2

Invention Disclosure

Rachata Ausavarungnirun, Gabriel Loh**.** Batcher and Batch Scheduler for a Simplistic Memory Request Scheduling. (Filed while at AMD)

Consultancy

Ditto Public Company Limited July 2023 – Now

- Tech lead for the Thailand National Zoo Project
  - Lead the datacenter design team
    - Generate high-level design based on demand from the user applications
    - Perform system design based on the entire software stack for the zoo (including software for zoo customers and animal management)
    - Work with hardware vendors, system integrators, and the construction team on the deployment of the datacenter
    - Ensure that the entire stack is reliable and resilient to faults
  - Lead the network infrastructure team
    - Generate network and system requirements based on client's demand
    - Coordinate network requirements and work with network and system integrator for the entire zoo network infrastructure
    - Work with the construction team on the network infrastructure
  - Assist the zoo project manager to ensure that the design is under the budget (multi-million dollars project)

PTT Exploration and Production (PTT-EP) July 2021 **–** Now
with The Gang Technology

- Tech Lead for the Sight SolutionS HPC project
  - Lead the design of an end-to-end GUI with PBS integration for HPC users
  - Aid the design of PTT-EP HPC infrastructure, including file system and network based on multiple teams' requirements
  - Accelerate and fine-tune oil and gas workloads' performance
- Tech Consultant for the INSURE project
  - Design a scalable authentication and management system for all internal files within PTTEP's storage servers
- Academic Coach for the Digital Citizen Bootcamp
  - Providing technical feedback for 4 teams in the bootcamp

Selected Top-tier Publications

Konstantinos Kanellopoulos, Rahul Bera, Kosta Stojiljkovic. Nisa Bostanci, Can Firtina, **Rachata Ausavarungnirun,** Rakesh Kumar, Nastaran Hajinazar, Mohammad Sadrosadati, Nandita Vijaykumar, Onur Mutlu *"Utopia: Efficient Address Translation using Hybrid Virtual-to-Physical Address Mapping"*. To appear in MICRO 2023.

Amel Fatima, Sihang Liu, Korakit Seemakhupt, **Rachata Ausavarungnirun**, and Samira Khan "*vPIM: Efficient Virtual Address Translation for Scalable Processing-in-Memory Architectures",* DAC 2023.

Changlong Li, Yu Liang, **Rachata Ausavarungnirun**, Zongwei Zhu, Liang Shi, Chuan Jason Xue **"***ICE: Collaborating Memory and Process Management for User Experience on Resource-limited Mobile Devices",* EuroSys 2023.

Xing Li, **Rachata Ausavarungnirun**, Xiao Liu, Xueyuan Liu, Xuan Zhang, Heng Lu, Zhuoran Song, Naifeng Jing and Xiaoyao Liang, "*Gzippo: Highly-compact Processing-In-Memory Graph Accelerator Alleviating Sparsity and Redundancy"*, ICCAD 2022.

Sangjin Choi, Taeksoo Kim, Jinwoo Jeong, **Rachata Ausavarungnirun**, Myeongjae Jeon, Youngjin Kwon, Jeongseob Ahn, "*Memory Harvesting in Multi-GPU Systems with Hierarchical Unified Virtual Memory"*, USENIX ATC 2022.

Nika Mansouri Ghiasi, Jisung Park, Harun Mustafa, Jeremie Kim, Arvid Gollwitzer, Ataberk Olgun, Haiyu Mao, Can Firtina, Damla Senol Cali, Nour Almadhoun Alserr, **Rachata Ausavarungnirun**, Nandita Vijaykumar, Mohammed Alser, Onur Mutlu, "*GenStore: An In-storage Processing System for Genome Sequence Analysis"*, ASPLOS 2022

Yu Liang, Riwei Pan, Tianyu Ren, and Yufei Cui, **Rachata Ausavarungnirun**, Xianzhang Chen, Changlong Li, Tei-Wei Kuo, Chun Jason Xue, "*CacheSifter: Sifting Cache Files for Boosted Mobile Performance and Lifetime"*, FAST 2022.

Xuanyi Li, Chen Li, Yang Guo, **Rachata Ausavarungnirun**, "*Improving Inter-kernel Data Reuse With CTA-Page Coordination in GPGPU"*, ICCAD 2021.

Maciej Besta, Raghavendra Kanakagiri, Grzegorz Kwasniewski, **Rachata Ausavarungnirun**, Jakub Beránek, Konstantinos Kanellopoulos, Kacper Janda, Zur Vonarburg-Shmaria, Lukas Gianinazzi, Ioana Stefan, Juan Gómez Luna, Jakub Golinowski, Marcin Copik, Lukas Kapp-Schwoerer, Salvatore Di Girolamo, Nils Blach, Marek Konieczny, Onur Mutlu, Torsten Hoefler, *"SISA: Set-centric instruction set architecture for graph mining on processing-in-memory systems"*, MICRO 2021**.**

Rachata Ausavarungnirun, Timothy Merrifield, Jayneel Gandhi, Christopher J Rossbach, "*PRISM: Architectural Support for Variable-granularity Memory Metadata"*, PACT 2020.

Damla Senol Cali, Gurpreet S. Kalsi, Zülal Bingöl, Lavanya Subramanian, CanFirtina, Jeremie Kim, **Rachata Ausavarungnirun**, Mohammed Alser, Anant Nori, Juan Gómez Luna, Amirali Boroumand, Allison Scibisz, Sreenivas Subramoney, Can Alkan, Saugata Ghose, Onur Mutlu. "*GenASM: A Low-Power, Memory-Efficient Approximate String Matching Acceleration Framework for Genome Sequence Analysis"*, MICRO 2020

Yu Liang, Jinheng Li, Xianzhang Chen, **Rachata Ausavarungnirun**, Riwei Pan, Tei-Wei Kuo, Chun Jason Xue. "*Differentiating Cache Files for Fine-grain Management to Improve Mobile Performance and Lifetime*", HotStorage 2020

Yu Liang, Jinheng Li, **Rachata Ausavarungnirun,** Riwei Pan, Liang Shi, Tei-Wei Kuo, Chun Jason Xue. "*Acclaim: Adaptive Memory Reclaim to Improve User Experience in Android Systems*", USENIX ATC 2020.

Nastaran Hajinazar, Pratyush Patel, Minesh Patel, Konstantinos Kanellopoulos, Saugata Ghose, **Rachata Ausavarungnirun**, Geraldo Francisco de Oliveira Junior, Jonathan Appavoo, Vivek Seshadri, Onur Mutlu. "*The Virtual Block Interface (VBI): A Flexible Alternative to Conventional Virtual Memory Frameworks"*. ISCA 2020.

Yu Liang, **Rachata Ausavarungnirun**, Tei-Wei Kuo, Chun Jason Xue, "*Uncovering Critical Data in Launching Mobile Applications to Improve User Experience at Low Cost",* FAST, 2020.

Xiao Liu, David Roberts, **Rachata Ausavarungnirun**, Onur Mutlu, Jishen Zhao. **"***Binary Star: Coordinated Reliability in Heterogeneous Memory Systems for High Performance and Scalability"*. MICRO 2019.

Amirali Boroumand, Saugata Ghose, Minesh Patel, **Rachata Ausavarungnirun**, Hasan Hassan, Brandon Lucia, Kevin Hsieh, Nastaran Hajinazar, Krishna T. Malladi, Hongzhong Zheng, Onur Mutlu. "*CoNDA: Enabling Efficient Near-Data Accelerator Communication by Optimizing Data Movement".* ISCA 2019.
Chen Li, **Rachata Ausavarungnirun**, Christopher J. Rossbach, Youtao Zhang, Onur Mutlu, Yang Guo, Jun Yang, "*A Framework for Memory Oversubscription Management in Graphics Processing Units",* ASPLOS 2019.

Rachata Ausavarungnirun, Vance Miller, Joshua Landgraf, Saugata Ghose, Jayneel Gandhi, Adwait Jog, Christopher J. Rossbach, and Onur Mutlu, "*MASK: Redesigning the GPU Memory Hierarchy to Support Multi-Application Concurrency",* ASPLOS 2018.

Amirali Boroumand, Saugata Ghose, Youngsok Kim, **Rachata Ausavarungnirun**, Eric Shiu, Rahul Thakur, Daehyun Kim, Aki Kuusela, Allan Knies, Parthasarathy Ranganathan, Onur Mutlu, *"Google Workloads for Consumer Devices: Mitigating Data Movement Bottlenecks", ASPLOS 2018*

Maciej Besta, Syed Minhaj Hassan, Sudhakar Yalamanchili, **Rachata Ausavarungnirun**, Onur Mutlu, Torsten Hoefler, "*Slim NoC: A Low-Diameter On-Chip Network Topology for High Energy Efficiency and Scalability",* ASPLOS 2018.

Mohammad Sadrosadati, Amirhossein Mirhosseini, Seyed Borna Ehsani, Hamid Sarbazi-Azad, Mario Drumond, Babak Falsafi, **Rachata Ausavarungnirun**, Onur Mutlu, "*LTRF: Enabling High-Capacity Register Files for GPUs via Hardware/Software Cooperative Register Prefetching",* ASPLOS 2018.

Rachata Ausavarungnirun, Joshua Landgraf, Vance Miller, Saugata Ghose, Jayneel Gandhi, Christopher J. Rossbach, and Onur Mutlu, "*Mosaic: A GPU Memory Manager with Application-Transparent Support for Multiple Page Sizes",* MICRO 2017.

Donghyuk Lee, Samira Khan, Lavanya Subramanian, Saugata Ghose, **Rachata Ausavarungnirun**, Gennady Pekhimenko, Vivek Seshadri, and Onur Mutlu, *"Design-Induced Latency Variation in Modern DRAM Chips: Characterization, Analysis, and Latency Reduction Mechanisms",* SIGMETRIC 2017.

Onur Kayıran, Adwait Jog, Ashutosh Pattnaik, **Rachata Ausavarungnirun**, Xulong Tang, Mahmut T. Kandemir, Gabriel H. Loh, Onur Mutlu and Chita R. Das, "*µC-States: Fine-grained GPU Datapath Power Management"*. PACT 2016.

Yang Li, Di Wang, Saugata Ghose, Jie Liu, Sriram Govindan, Sean James, Eric Peterson, John Siegler, **Rachata Ausavarungnirun**, and Onur Mutlu, "*SizeCap: Efficiently Handling Power Surges in Fuel Cell Powered Data Centers"*. HPCA 2016.

Rachata Ausavarungnirun, Chris Fallin, Xiangyao Yu, Kevin Chang, Greg Nazario, Reetuparna Das, Gabriel Loh, and Onur Mutlu, "*A Case for Hierarchical Rings with Deflection Routing: An Energy-Efficient On-Chip Communication Substrate"*. PARCO 2016.

Rachata Ausavarungnirun, Saugata Ghose, Onur Kayiran, Gabriel H. Loh, Chita R. Das, Mahmut T. Kandemir, Onur Mutlu, "*Exploiting Inter-Warp Heterogeneity to Improve GPGPU Performance"*. PACT 2015**.**

Donghyuk Lee, Lavanya Subramanian, **Rachata Ausavarungnirun**, Jongmoo Choi, Onur Mutlu, "*Decoupled Direct Memory Access: Isolating CPU and IO Traffic by Leveraging a Dual-Port DRAM"*. PACT 2015.

Mohammad Fattah, Antti Airola, **Rachata Ausavarungnirun**, Nima Mirzaei, Pasi Liljeberg, Juha Plosila, Siamak Mohammadi, Tapio Pahikkala, Onur Mutlu and Hannu Tenhunen, "*A Low-Overhead, Fully-Distributed, Guaranteed-Delivery Routing Algorithm for Faulty Network-on-Chips"*. NoCs 2015.

Nandita Vijaykumar, Gennady Pekhimenko, Adwait Jog, Abhishek Bhowmick, **Rachata Ausavarungnirun**, Onur Mutlu, Chita R. Das, Mahmut T. Kandemir, Todd C. Mowry *"A Case for Core-Assisted Bottleneck Acceleration in GPUs: Enabling Efficient Data Compression"*. ISCA 2015.

Onur Kayiran, Nachiappan Chidambaram, Adwait Jog, **Rachata Ausavarungnirun**, Mahmut T. Kandemir, Gabriel Loh, Onur Mutlu, Chita R. Das, "*Managing GPU Concurrency in Heterogeneous Architectures",* MICRO 2014.

Rachata Ausavarungnirun, Chris Fallin, Xiangyao Yu, Kevin Chang, Greg Nazario, Reetuparna Das, Gabriel Loh, Onur Mutlu, "*Design and Evaluation of Hierarchical Rings with Deflection Routing"*, SBAC-PAD 2014.

ASST. PROF. RACHATA AUSAVARUNGNIRUN (HAVE GREEN CARD)

Thai-German Graduate School, KMUTNB, Bangkok, Thailand
r.ausavarungnirun@gmail.com

Vivek Seshadri, Yoongu Kim, Chris Fallin, Donghyuk Lee, **Rachata Ausavarungnirun**, Gennady Pekhimenko, Yixin Luo, Onur Mutlu, Phillip B. Gibbons, Michael A. Kozuch. Todd C. Mowry "*RowClone: Fast and Efficient In-DRAM Copy and Initialization of Bulk Data".* MICRO 2013.

Reetuparna Das, **Rachata Ausavarungnirun**, Onur Mutlu, Akhilesh Kumar, Mani Azimi. "*Application-to-Core Mapping Policies to Reduce Memory Interference in Multi-Core Systems"*. HPCA 2013.

HanBin Yoon, Justin Meza, **Rachata Ausavarungnirun**, Racheal A. Harding, Onur Mutlu. *"Row Buffer Locality Aware Caching Policies for Hybrid Memories",* ICCD 2012*.* ***Best Paper Award (in Computer Systems and Applications track).***

Rachata Ausavarungnirun, Kevin Chang, Lavanya Subramanian, Gabriel Loh, Onur Mutlu. *"Staged Memory Scheduling: Achieving High Performance and Scalability in Heterogeneous Systems".* ISCA 2012.

Chris Fallin, Greg Nazario, Xiangyao Yu, Kevin Chang, **Rachata Ausavarungnirun**, Onur Mutlu. *"MinBD: A Minimally-Buffered Deflection Router Approaching Conventional Buffered-Router Performance"*. NoCs 2012. ***One of the five papers nominated for the Best Paper Award by the Program Committee.***

Honors

The **fourth person in Asia** to be indicted to the **ASPLOS Hall of Fame**	2022
Bertucci Fellowship in Engineering	2014
Best paper award in ICCD conference	2012
Carnegie Institute of Technology Honors Research Program	2010
Carnegie Mellon University Honors	2010
Royal Thai Government Scholarship – 11-year full scholarship	2005

Professional Service

Program Committee:	ASPLOS 2025, HPCA 2024, MICRO 2023, IIWSC 2022, ISPASS 2022, MICRO 2021, ISPASS 2021, MICRO 2020, GPGPU 2020, GPGPU 2019
Finance Chair:	ISPASS 2022
External Review Committee:	ASPLOS 2023, MICRO 2022, ISCA 2022, ISCA 2021, HPCA 2021, ISCA 2020, ISCA 2019

Reviewer: ACM TACO, IEEE CAL, IEEE TVSLI, IEEE, TPDS, IEEE TC, IEEE TCAS