

TONY NOWATZKI

310-825-8807 (Cell: 612-269-6036) • tjn@cs.ucla.edu
404 Westwood Plaza, Eng. VI Room 468B • Los Angeles, CA 90095

RESEARCH INTERESTS

Computer Architecture – ISA Specialization / Accelerators / Near-data Processing
Compilers – Dynamic Compilation / Compilers for Spatial & Reconfigurable / Domain-Specific
Optimization – Design-Space Exploration / Instruction Scheduling / Constraint & ML-based

EDUCATION

University of Wisconsin, Madison (4.00) • Ph.D. in Computer Sciences (Fall 2016)

University of Wisconsin, Madison (4.00) • MS in Computer Sciences (Spring 2011)

University of Minnesota, Twin Cities (3.95) • B.S. in Comp. Sciences & Comp. Engineering (Spring 2009)

AWARDS AND HONORS

- IEEE Micro Top Picks, 2022, 2020, 2017, 2016 (2021 – Honorable mention)
- MICRO 2022 Best Paper Nomination
- HPCA 2021 Runner-up to Best Paper
- CACM Research Highlights, 2019
- NSF CAREER Award, 2018
- Best of CAL, 2015
- Google PhD Fellowship in Computer Architecture, 2014
- Distinguished Paper Award, PLDI 2013
- Publication Nominated by SIGPLAN for a CACM Research Highlights, 2013

ACADEMIC EXPERIENCE

University of California Los Angeles, PolyArch Research Group (January 2017–Present)

Assistant Professor

- Formed the PolyArch research group to study design of heterogeneous programmable processors and their compilers. Major research vectors:
- **Broadening the Scope of Programmable Accelerators** *Problem:* Existing vector-style processors and accelerators stumble on irregular workloads (irregular control, memory, parallelism) *Approach:* Find common idioms between different domains, and develop classes of programmable accelerators with flexible primitives.
- **Automating Accelerator Design** *Problem:* Developing a SW/HW stack for any given accelerator is a challenging multi-year effort. *Approach:* Create common representations for modular accelerator primitives, along with compiler algorithms to drive design-space search.
- **Bridging the Accelerator/Von Neumann Gap** *Problem:* CPU architectures struggle due to sequential ISAs with implicit dependences and no high-level information. *Approach:* Develop new ISAs that express rich computation and memory semantics to hardware, and take advantage transparent use of specialized hardware and near-data processing through the memory hierarchy.

University of Wisconsin, Vertical Research Group (June 2010–December 2016)

Research Assistant

- Design and evaluation of modular general purpose processor architecture, including several in-core dataflow architectures.
- Modeling techniques for fundamentally understanding and designing programmable accelerators.
- Mathematical optimization-based compiler for general spatial scheduling using Integer Linear Programming.

University of Wisconsin, Department of Computer Sciences (Fall 2009–Spring 2011)
Teaching Assistant

- Worked with professors to develop curriculum for two courses in computer architecture.
- Gained leadership experience teaching in a classroom and in one-on-one environments.

University of Minnesota, Digital Technology Center (2005–2009)

Research Assistant

- Worked with a team to organize and run large scale physics simulations on clusters and super computers.
- Lead developer on C# application for visualizing 3D volumetric data, which leveraged a cluster of rendering nodes for real-time data exploration and cinematic capture/playback.

TEACHING

CS33: Introduction to Computer Organization (2017–Present)

- This course demystifies computer systems, covering the basics of computer architecture, computer organization, operating systems and concurrency.
- Uses four interactive and auto-graded projects designed to provoke interest among newcomers. My offering targets those coming from fields other than CS.

CS251a: Introduction to Computer Organization (2018–Present)

- Explores the art and science behind creating architecture abstractions that enable efficient use of hardware without sacrificing generality.
- We show how to use intrinsic properties of applications to reason the tradeoffs of hardware components to meet cost and performance goals.

CS259: Learning Machines (2018–Present)

- This course explores – from a computer architecture perspective – the principles of hardware/software codesign for machine learning.
- We discuss both accelerator innovations for ML algorithms, including parallelization and data-orchestration, and also explore how ML can be used to improve general purpose processors.

CS259: Heterogeneity and the Specialization Spectrum (2017)

- This class explores the benefits and drawbacks of specialization, as well as the challenges in architecture design, compilation and programming models.

INDUSTRY EXPERIENCE

Simple Machines Inc. (2017 – 2021)

Consultant

- Compiler and ISA development for SMI's composable compute platform for accelerating AI and ML environments.

Qualcomm, Qualcomm Research Silicon Valley (Summer 2011 & 2012)

Internship

- Implementing and evaluating next-generation dynamic compilers.

CONFERENCE PUBLICATIONS

- Zhengrong Wang, Christopher Liu, Tony Nowatzki. Infinity Stream: Portable and Programmer-Friendly In-/Near-Memory Fusion. *ASPLOS 2023*
- Sihao Liu, Jian Weng, Dylan Kupsh, Atefeh Sohrabizadeh, Zhengrong Wang, Licheng Guo, Jiuyang Liu, Maxim Zhulin, Lucheng Zhang, Jason Cong, Tony Nowatzki OverGen: Improving FPGA Usability through Domain-specific Overlay Generation *MICRO 2022 – Best Paper Nominee*

- Karthikeyan Sankaralingam, Tony Nowatzki, Vinay Gangadhar, Preyas Shah, Michael Davies, Poly Palamuttam, Jitendra Khare, Maghawan Punde, Deepak Vijay, Ziliang Guo, William Galliher, Vinay Thiruvengadam, Alex Tan. The Mozart Reuse Exposed Dataflow Processor for AI and Beyond. *ISCA* 2022
- Vidushi Dadu, Sihao Liu, Tony Nowatzki. TaskStream: Accelerating Task-Parallel Workloads by Recovering Program Structure. *ASPLOS* 2022
- Zhengrong Wang, Jian weng, Sihao Liu, Tony Nowatzki. Near-Stream Computing: General and Transparent Near-Cache Acceleration. *HPCA* 2022
- Vidushi Dadu, Sihao Liu, Tony Nowatzki. PolyGraph: Exposing the Value of Flexibility for Graph Processing Accelerators. *ISCA* 2021 – **IEEE Micro Top Picks 2022**
- Zhengrong Wang, Jian Weng, Jason Lowe-Power, Jayesh Gaur, Tony Nowatzki. Stream Floating: Enabling Proactive and Decentralized Cache Optimizations. *HPCA* 2021 – **Best Paper Runner-up**
- Jian Weng, Animesh Jain, Jie Wang, Leyuan Wang, Yida Wang, Tony Nowatzki. UNIT: Unifying Tensorized Instruction Compilation. *CGO* 2021
- Jian Weng, Sihao Liu, Vidushi Dadu, Zhengrong Wang, Preyas Shah, Tony Nowatzki. DSAGEN: Synthesizing Programmable Spatial Accelerators. *ISCA* 2020
- Jian Weng, Sihao Liu, Zhengrong Wang, Vidushi Dadu, Tony Nowatzki. A Hybrid Systolic-Dataflow Architecture for Inductive Matrix Algorithms. *HPCA* 2020
- Vidushi Dadu, Jian Weng, Sihao Liu, Tony Nowatzki. Towards General Purpose Acceleration by Exploiting Common Data-Dependence Forms. *MICRO* 2019 – **IEEE Micro Top Picks 2020**
- Amirali Sharifian, Reza Hojabr, Navid Rahimi, Sihao Liu, Apala Guha, Tony Nowatzki, Arrvindh Shriraman. μ IR -An intermediate representation for transforming and optimizing the microarchitecture of application Accelerators. *MICRO* 2019
- Zhengrong Wang, Tony Nowatzki. Stream-based Memory Access Specialization for General Purpose Processors. *ISCA* 2019
- Tony Nowatzki, Newsha Ardalani, Karthikeyan Sankaralingam, Jian Weng. Hybrid Optimization/Heuristic Instruction Scheduling for Programmable Accelerator Codesign. *PACT* 2018.
- Tony Nowatzki, Vinay Gangadhar, Newsha Ardalani, Karthikeyan Sankaralingam. Stream-Dataflow Acceleration *ISCA* 2017
- Tony Nowatzki, Vinay Gangadhar, Karthikeyan Sankaralingam, Greg Wright. Pushing the Limits of Accelerator Efficiency While Retaining General-Purpose Programmability. *HPCA* 2016 – **IEEE Micro Top Picks 2017**
- Tony Nowatzki, Karthikeyan Sankaralingam. A Framework and Analysis of Behavior Specialized Accelerators. *ASPLOS* 2016
- Matthew Watkins, Tony Nowatzki, Anthony Carno. Software Transparent Dynamic Binary Translation for Coarse-Grain Reconfigurable Architectures. *HPCA* 2016
- Tony Nowatzki, Vinay Gangadhar, Karthikeyan Sankaralingam. Exploring the Potential of Heterogeneous Von Neumann/Dataflow Execution Models. *ISCA* 2015 – **IEEE Micro Top Picks 2016**
- Chen-han Ho, Vinay Govindaraju, Tony Nowatzki, R. Nagaraju, Z. Marzec, P. Agarwal, C. Frericks, R. Cofell, Karthikeyan Sankaralingam. Performance evaluation of a DySER FPGA prototype system spanning the compiler, microarchitecture, and hardware implementation. *ISPASS* 2015
- Tony Nowatzki, M. Sartin-Tarm, L. De Carli, Karthikeyan Sankaralingam, C. Estan, B. Robotmili. A General Constraint-centric Scheduling Framework for Spatial Architectures. *PLDI* 2013 – **Distinguished Paper Award**
- Venkatraman Govindaraju, Tony Nowatzki, Karthikeyan Sankaralingam. Breaking SIMD Shackles with an Exposed Flexible Microarchitecture and the Access Execute PDG. *PACT* 2013
- Jesse Benson, Ryan Cofell, Chris Frericks, Chen-han Ho, Venkatraman Govindaraju, Tony Nowatzki, Karthikeyan Sankaralingam. Design Integration and Implementation of the DySER Hardware Accelerator into OpenSPARC. *HPCA* 2012

- Jian Weng, Sihao Liu, Dylan Kupsh, Tony Nowatzki Unifying Spatial Accelerator Compilation with Idiomatic and Modular Transformations. *IEEE Micro Special Issue on Compiling for Accelerators*, 2022
- Zhengrong Wang, Christopher Liu, Tony Nowatzki. Infinity Stream: Enabling Transparent and Automated In-Memory Computing. *CAL 2022*
- Vidushi Dadu, Sihao Liu, Tony Nowatzki. Systematically Understanding Graph Accelerator Dimensions and the Value of Hardware Flexibility. *IEEE Micro Top Picks in Computer Architecture 2022*
- Shail Dave, Riyadh Baghdadi, Tony Nowatzki, Sasikanth Avancha, Aviral Shrivastava, Baoxin Li. Hardware Acceleration of Sparse and Irregular Tensor Computations of ML Models: A Survey and Insights. *PIEE 2021*
- Vidushi Dadu, Jian Weng, Sihao Liu, Tony Nowatzki. Towards General Purpose Acceleration: Finding Structure in Irregularity. *IEEE Micro Top Picks in Computer Architecture 2020*
- Tony Nowatzki, Vinay Gangadhar, Karthikeyan Sankaralingam. Heterogeneous Von Neumann/Dataflow Microprocessors. *Communications of the ACM 2019*
- Jian Weng, Sihao Liu, Vidushi Dadu, Tony Nowatzki. DAEGEN: A Modular Compiler for Exploring Decoupled Spatial Accelerators. *CAL 2019*
- Gagan Gupta, Tony Nowatzki, Vinay Gangadhar, and Karthikeyan Sankaralingam. Kickstarting Semiconductor Innovation with Open Source Hardware. *IEEE Computer 2017*
- Tony Nowatzki, Venkatraman Gangadhar, Karthikeyan Sankaralingam, Greg Wright. Domain Specialization is Generally Unnecessary for Accelerators. *IEEE Micro Top Picks in Computer Architecture 2017*
- Tony Nowatzki, Jai Menon, Chen-han Ho, Karthikeyan Sankaralingam. Architectural Simulators Considered Harmful. *IEEE Micro, 2015*
- Amir Yazdanbakhsh, Raghuraman Balasubramanian, Tony Nowatzki, Karthikeyan Sankaralingam. Comprehensive Circuit Failure Prediction and Detection for Logic and SRAM using Virtual Aging, Sampled Redundancy, and Asymmetric Checkers. *IEEE Micro 2015*
- Tony Nowatzki, Venkatraman Govindaraju, Karthikeyan Sankaralingam. A Graph-Based Program Representation for Analyzing Hardware Specialization Approaches. *CAL 2015 – Best of CAL Award*
- Tony Nowatzki, Michael Sartin-Tarm, Lorenzo De Carli, Karthikeyan Sankaralingam, Cristian Estan, Behnam Robotmili. A Scheduling Framework for Spatial Architectures Across Multiple Constraint-solving Theories. *TOPLAS 2014*
- Tony Nowatzki Jai Menon, Chen-han Ho, Karthikeyan Sankaralingam. gem5, GPGPUSim, McPAT, GPUWattch, "Your favorite simulator here" Considered Harmful. *WDDD 2014*
- Michael Sartin-Tarm, Tony Nowatzki, Lorenzo De Carli, Karthikeyan Sankaralingam, Cristian Estan. Constraint centric scheduling guide. *SIGARCH Comput. Archit. News 2013*
- Venkatraman Govindaraju, Chen-han Ho, Tony Nowatzki, J. Chhugani, N. Satish, Karthikeyan Sankaralingam, C. Kim. DySER: Unifying Functionality and Parallelism Specialization for Energy Efficient Computing. *IEEE Micro 2012*
- P. Woodward, J. Jayaraj, P. Lin, P. Yew, M. Knox, J. Greensky, T. Nowatzki, K. Stoffels. Boosting the performance of computational fluid dynamics codes for interactive supercomputing. *Int. Conf. on Membrane Computing 2010*
- P. Woodward, F. Herwig, D. Porter, T. Fuchs, T. Nowatzki. Nuclear burning and mixing in the first stars: Entrainment at a convective boundary using the PPB advection scheme. *AIP 2008*

BOOKS

- Tony Nowatzki, Michael Ferris, Karthikeyan Sankaralingam, Cristian Estan, Nilay Vaish, David Wood. Optimization and Mathematical Modeling in Computer Architecture. *Synthesis Lectures on Computer Architecture*, September 2013.

PATENTS/DISCLOSURES

- K. Sankaralingam, T. Nowatzki, V. Gangadhar. Computer architecture with fixed program dataflow elements and stream processor, Issued October 19, 2021

- K. Sankaralingam, T. Nowatzki, V. Gangadhar, . Shah, N. Ardalani. Systems and methods for stream-dataflow acceleration wherein a delay is implemented so as to equalize arrival times of data packets at a destination functional unit, Issued June 29, 2021
- K. Sankaralingam, Y. Li, V. Gangadhar, T. Nowatzki, Accelerating parallel processing of data in a recurrent neural network, Issued June 22, 2021
- K. Sankaralingam, V. Gangadhar, T. Nowatzki, Y. Li. Method, computer program product, and apparatus for acceleration of simultaneous access to shared data, Issued March 31 2021
- T. Nowatzki, V. Gangadhar, K Sankaralingam. Computer with Hybrid Von-Neumann/Dataflow Execution Architecture, US P150319US01, Issued Feb 2019
- A. Yazdanbakhsh, R. Balasubramanian, T. Nowatzki, K. Sankaralingam. Computer System Predicting Memory Failure, US P150070US01, Issued March 2016

WORKSHOPS/TUTORIALS ORGANIZED

- Undergrad Architecture Mentoring (uArch) Workshop, ISCA 2021-2023
- OverGen: Spatial Architecture Synthesis for ASICs and FPGA Overlays, MICRO 2022
- DSAGEN: Democratizing Spatial Accelerator Research, MICRO 2020
- Renegotiating the Levels of Abstraction for the Post Moore's Law Era, ARM Research Summit, 2019

PROGRAM COMMITTEE AND REVIEWING SERVICE

- International Symposium on Computer Architecture (ISCA), Reviewer, 2023
- Young Architect Workshop (YArch), Program Committee Member, 2023
- Architectural Support for Programming Languages and Operating Systems (ASPLOS), Reviewer, 2023 (spring, summer, and fall committees)
- International Symposium on Workload Characterization, Artifact Evaluation Co-chair, 2022
- International Symposium on Computer Architecture (ISCA), External Reviewer, 2022
- Young Architect Workshop (YArch), Program Committee Member, 2022
- International Symposium on Microarchitecture (MICRO), Program Committee Member, 2021
- Transactions on Parallel and Distributed Systems (TPDS), Reviewer, 2021
- International Conference on Supercomputing (ICS), Program Committee Member, 2021
- Young Architect Workshop (YArch), Program Committee Member, 2021
- Symposium on High-Performance Computer Architecture (HPCA), Program Committee Member, 2021
- Architectural Support for Programming Languages and Operating Systems (ASPLOS), External Reviewer, 2021
- International Symposium on Microarchitecture (MICRO), Program Committee Member, 2020
- International Symposium on Computer Architecture (ISCA), Program Committee Member, 2020
- Young Architect Workshop (YArch), Program Committee Member, 2020
- Design Automation Conference (DAC), Program Committee Member, 2019
- IEEE International Conference on Computer Design (ICCD), Program Committee Member, 2019
- 52nd IEEE/ACM International Symposium on Microarchitecture (MICRO), External Reviewer, 2019
- International Symposium on Computer Architecture (ISCA), External Reviewer, 2019
- Symposium on High-Performance Computer Architecture (HPCA), External Reviewer, 2019
- International Symposium on Memory Management (ISMM), Program Committee Member, 2019
- International Workshop on Exploitation of High Performance Heterogeneous Architectures and Accelerators (WEHA) Program Committee Member, 2019
- Transactions on Computers (TC), Reviewer, 2019
- Computer Architecture Letters (CAL), Reviewer, 2019
- Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD), Reviewer, 2019
- Transactions on Architecture and Code Optimization (TACO), Reviewer, 2019

- IEEE International Conference on Computer Design (ICCD), Program Committee Member, 2018
- International Workshop on Exploitation of High Performance Heterogeneous Architectures and Accelerators (WEHA), Program Committee Member, 2018
- IEEE Journal on Emerging and Selected Topics in Circuits and Systems, Reviewer, 2018
- IEEE Computer Architecture Letters, Reviewer, 2018
- Transactions on Architecture and Code Optimization (TACO), Reviewer, 2018
- Transactions on Computers, Reviewer, 2018
- Micro 2017, Program Committee Member, 2017
- Reviewer for Transactions on Architecture and Code Optimization, 2017
- Reviewer for Transactions on Computers, 2017
- Reviewer for Computer Architecture Letters, 2017

PHD COMMITTEE SERVICE

- Khanh Truong D. Nguyen, CS
- Nazanin Farahpour, CS
- Feng Shi, CS
- Aayush Jain, CS
- Renju Liu, CS
- Peipei Zhou, CS
- Young-Kyu Choi, CS
- Yuchen Hao, CS
- Yuze Chi, CS
- Jie Wang, CS
- Sandeep Singh, CS
- Vikranth Jeyakumar, EE
- Uneeb Yaqub Rathore, EE
- Saptadeep Pal, EE
- Irina Alam, EE
- Wojciech Romaszkan, EE