Weiwei Chen

Homepage: weiweichen.github.io   **Email:** weiwei.chen.uci@gmail.com

Research Interests

- Compilers and Frameworks for Machine Learning, High-Performance Computing, and Big Data
- Heterogeneous Parallel Programming

Education

Ph.D. Electrical and Computer Engineering - University of California, Irvine	2013

Dissertation: Out-of-Order Parallel Discrete Event Simulation for Electronic System Level Design
Outstanding Dissertation Award, European Design and Automation Association

M.S. Computer Engineering - Shanghai Jiao Tong University	2007
B.S. Computer Science and Engineering - Shanghai Jiao Tong University	2004

.

Experience

Senior Staff Engineer - Modular	March 2025 - Present
Staff Engineer - Modular	May 2023 - Feb 2025

Building next-generation AI infrastructure and develop platform

- Compiler for the Mojo programming language.

Senior Principal Engineer - SambaNova Systems	May 2021 - May 2023
Principal Engineer - SambaNova Systems	Oct 2018 - May 2021

Building Compilers for Coarse-Grain Reconfigurable Dataflow Units (RDUs)

- **Arc** - An MLIR-based DSL and Compiler for stitching high performance tensor operation templates on RDUs, including Layout and Control Analysis, Buffer and Anti-hang Peephole Optimizations, Template Selection and Lowering, etc. One of the first two engineers brought up this layer.
- **RAIL** - An MLIR-based DSL and Compiler for programming the compute and memory components on RDUs. Founding member of the project and main contributor to the compute component, including syntax design, register allocation, context splitting, latency analysis, etc.
- **Nova** - PyTorch JIT for RDU, including adding dynamic dispatch key for RDU, RDU compiler integration using lazy tensor, hot/cold caching mechanisms, performance annotation for auto-differentiated graphs, heterogeneous backend enabling for CPU+GPU+RDU, etc. One of the first two engineers bringing up this layer.
- **Assembler** for multi-generation RDUs, founding member of the project to design and implement an in-house data structure for the memory component and assembler in general.
- Various compiler infrastructure tasks, including LLVM/MLIR version upgrading; architecting and migrating one compiler layer from Python to MLIR in C++; migrating RDU backend from Python to C++; cmake structuring and maintenance; ML model compilation debugging support, etc
- Various compiler PoC projects, including a Spatial-like RDU compiler in Scala; JAX to RDU lowering; new DSL for RDU programming at higher abstraction level than RAIL, etc.

Founding Software Engineer - BigStream Solutions	Feb 2016 - Oct 2018

First Engineer hired to lead software research and development efforts on compilers, native C++ acceleration libraries, tool-chains for FPGA acceleration, and big data system architecture.

- A dataflow compiler in Scala for Spark SQL query to intermediate representations with three backend for native CPU, FPGA, and RISCV code generation, and a planner for optimized SW/HW accelerator partitioning.
- Native C++ runtime and compiler support for accelerating user-defined functions(UDF) in Scala.
- A C++ native acceleration library with templatized SQL operations, cluster data source support, (HDFS, amazon S3, Microsoft WASB), various input format support (json, avro, parquet, csv) and TensorFlow native integration for data pipelines.
- A Clang-LLVM based high-level synthesis compiler for timing scheduling and code generation for FPGA.
- Various engineering support, including on-premise Hadoop Cluster setup, network performance measurement, etc.

Senior Engineer - Qualcomm Research Silicon Valley   Oct 2013 - Feb 2016

Member of the Pervasive Concurrency team for Qualcomm Symphony System Manager SDK, a task-based parallel programming patterns and runtime for heterogeneous multi-core platforms.

- Heterogeneous parallel pipeline programming pattern API and internal scheduling, task and dataflow API infrastructures, parallelize Android native computational photography and enterprise compression applications using task-based parallel patterns, power and performance evaluation for native parallel applications, etc.
- Compiler frontend analysis and backend code generation for coarse-grain auto-parallelization base on polyhedral optimizations in LLVM.

Software Develop Engineer Intern - Microsoft   June 2011 - Sept 2011

Windows Core Security and Identity Public Key Infrastructure Team

- A Windows store application for secure banking on Windows 8.

Patents (selected)

- Iteration Synchronization Construct for Parallel Pipelines (granted), US 15/191,266
- Systems And Methods For Accelerating Data Operations By Utilizing Native Memory Management (filed),US 62/775,533
- Systems and Methods for Accelerating Data Operations By Utilizing Dataflow Subgraph Templates (filed), US 16/898,048
- Compile Time Logic for Inserting a Buffer Between a Producer Operation Unit and a Consumer Operation Unit in a Dataflow Graph (granted), US 17/582,421
- Compile Time Logic for Detecting Streaming Compatible and Broadcast Compatible Data Access Patterns (filed), US 17/031,679
- Anti-Congestion Flow Control for Reconfigurable Processors (filed), US 16/890,841
- Partitioning Dataflow Operations for a Reconfigurable Computing System (filed), US 63/317,476

Publications (selected)

- Weiwei Chen, Xu Han, Rainer Doemer, "May-Happen-in-Parallel Analysis based on Segment Graphs for Safe ESL Models", in Proceedings of the *Design, Automation and Test in Europe Conference (DATE)*, Dresden, Germany, March 2014 **(Best Paper Award)**
- Weiwei Chen, Xu Han, Che-Wei Chang, Guantao Liu, Rainer Doemer, "Out-of-Order Parallel Discrete Event Simulation for Transaction Level Models", *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD)*, vol.33, no.12, pp.1859-1872, Dec 2014
- 3 Journal papers, 1 Book, 2 Book Chapters, 14 Conference papers, 4 Technical Reports, 4 Poster Presentations, see full list from here.

Awards and Honors (selected)

- 7 Qualstar Awards, Qualcomm Inc. 2014 - 2016
- Outstanding Dissertation Award, European Design and Automation Association (EDAA) 2014
- Best Paper Award, Design, Automation and Test Conference in Europe (DATE) 2014
- Pedagogical Fellowship, UC Irvine 2012-13
- Henry Samueli Endowed Fellowship, UC Irvine 2007
- National Scholarship for Academic Excellence, China 2006
- Exceptional Undergraduate Student Awards, SJTU
- People's Scholarship for Academic Excellence, SJTU 2000-2004

Skills and Hobbies

- C++ (native), Scala (product), Python (project), Rust (Advent of Code), Haskell (Advent of Code), Git, PyTorch, Vim, GDB, Valgrind, cmake, Latex ...
- Board member of San Francisco Bay Queer Contra Dance
- Fiddling, Reading, Hiking, Folk Dancing, Gardening, Karate, Sailing, Board Games

Working Status

- US Citizen