Gabriel Weisz, PhD

✉ gabeweisz@gmail.com
🖥 gabeweisz.com
in gabriel-weisz

Summary

- Experienced technology leader that has worked as an individual contributor, team lead, and manager at small and large companies
- Created systems for a variety of fields including neural network compilers and co-designed hardware accelerators, web and email infrastructure, and enterprise healthcare software
- Comfortable with all levels of the software stack from desktop, web, and mobile clients to back-end interfaces, databases, and services
- U.S. citizen located in Bethesda, MD

Experience

2024-Present   **Fellow Software Development Engineer**, *Advanced Micro Devices*.
- Followed the latest developments in AI models and techniques
- Identified the most critical and relevant developments and worked with team members to prototype and analyze them on AMD platforms
- Developed innovative new techniques that leverage the strength of AMD platforms for improved performance and higher quality and shared them with the AI community through publications and open-source contributions
- Communicated findings with AMD internal teams and the external AI community
- Identified strategic opportunities and partnerships for AMD in emerging AI technologies

2022-2024   **Director of Product Management**, *MangoBoost, Inc.*
- Worked to turn academic research prototypes into a Data Processing Unit product
- Determined which features and performance metrics to prioritize
- Advised the development team on architectural and engineering decisions
- Defined development procedures and best practices related to testing and product releases
- Developed pricing models and estimated the size of the potential market
- Collected market research on the industry and potential competitors

2021-2022   **Principal Hardware Engineering Manager**, *Microsoft*.

2019-2021   **Principal Hardware Engineer**, *Microsoft*.

2017-2019   **Senior Hardware Engineer**, *Microsoft*.
- Developed and managed a team of engineers developing Brainwave, an FPGA-accelerated neural network inference system used by the Bing search engine and M365
- Prototyped and spearheaded the development of a model compiler that ports neural network models created in TensorFlow, PyTorch, ONNX, and other neural network development frameworks to Brainwave
- Worked with internal customers to assist in porting models to Brainwave
- Developed RTL hardware for the FPGA-based Brainwave Neural Processing Unit
- Developed software for the Brainwave runtime system
- Validated and tested the correctness and end-to-end performance of deep neural network models running on Brainwave, including evaluating the effects of reduced-precision computations on inference accuracy
- Presented Brainwave at various conferences

| 2015-2017 | **Computer Scientist**, *University of Southern California, Information Sciences Institute*. |

- Research on hardware systems for high performance computing with a focus on FPGAs
- NASA-funded SpaceCubeX project supporting heterogeneous application development across FPGAs, DSPs, and microcontrollers
- DARPA-funded HAVoC project developing large, real-world benchmark applications for third-party hardware malware detection systems
- Investigated netlist analysis and obfuscation mechansims for security applications
- Prototyped, gathered data for, and collaborated on grant proposals
- Presented our research at program reviews

| 2010-2015 | **Graduate Research Assistant**, *Carnegie Mellon University, Department of Computer Science*. |

- Research at the Computer Architecture Lab at Carnegie Mellon University focused on making FPGA devices easier to use while maximizing performance
- CoRAM++: An FPGA application development framework providing a C-like interface to hardware-accelerated data structures and memory interfaces
- GraphGen: A graph processing framework for mapping vertex-centric graph applications to FPGAs
- C-to-CoRAM: A high-level synthesis framework built within the LLVM compiler that automatically mapped computation kernels and their DRAM accesses to FPGAs

| 1999-2010 | **Co-founder and Vice President of Technology**, *Salar, Inc*. |

- Bootstrapped the company from inception to millions of dollars in annual revenue, customers among the top hospitals in the country, and an award for Baltimore's best place to work in 2009
- Developed software and managed a team of software developers on projects including:
  - TeamNotes, Salar's flagship documentation and charge capture product, in use by over 25 hospitals to document over a million patient record per year
  - An embeddable version of the TeamNotes forms engine that was licensed to several customers
  - Tap, a Palm OS physician professional fee charge capture system
  - The Palm OS version of the Johns Hopkins Antibiotic Guide, a handheld clinical drug reference that was installed by over 150,000 people worldwide
  - A document authoring and peer-review system for reviewed medical references that publishes to the web, print, and mobile devices

| 1999 | **Software Engineer**, *Immersive Technologies, LLC*. |

Developed software in the company's platform for creating cinematic, photo-realistic images using beam tracing techniques and a novel special subdivision mechanism

| 1998–1999 | **Software Engineer**, *Jump.com*. |

Developed one of the first free web-based email, shared calendar, and contact management systems

Education

| PhD | **Carnegie Mellon University, Department of Computer Science**. |

Awarded December 2015. Advisor: James C. Hoe.

| MS | **Carnegie Mellon University, Department of Computer Science**. |

Awarded May 2013

| BS | **Cornell University, College of Engineering**, *Cum Laude*. |

Double major in computer science and electrical engineering. Awarded May 1999

Skills

Technical leadership: System architecture, design, and development of thick client applications, mobile applications, deep learning systems, hardware/software codesign, compilers, web applications, enterprise and distributed systems, security infrastructure, and databases

Project management: Requirements gathering, specification development, timeline creation, testing and issue tracking, revision control, configuration and release management, code reviews, agile methodologies and scrum

Programming languages: C, C++, C#, Java, JavaScript, SQL, Python, Perl, PHP, VB, Matlab, XML, HTML, Verilog, VHDL, Bluespec System Verilog

Development environments and technologies: Microsoft Visual Studio, Eclipse, gcc/make, .Net, MFC, ADO, COM, ODBC, DAO, ActiveX, OpenGL, DirectX

Databases: Microsoft SQL Server, Oracle, MySQL, Sqlite

Revision Control: Git, Azure Devops, subversion, Perforce, Visual Source Safe

Patents

2023 **Patent #11,556,762** Neural network processor based on application specific synthesis specialization parameters. Fowers, Ovtcharov, Chung, Massengill, Liu, & Weisz

2020 **Patent #10,795,678** Matrix vector multiplier with a vector register file comprising a multi-port memory. Fowers, Ovtcharov, Chung, Massengill, Liu, & Weisz

2012 **Patent #8,326,653** Method and apparatus for analyzing patient medical records. Gottlieb, Gottlieb, & Weisz

Selected Publications

ISCA 2018 Jeremy Fowers, Kalin Ovtcharov, Michael Papamichael, Todd Massengill, Ming Liu, Daniel Lo, Shlomi Alkalay, Michael Haselman, Logan Adams, Mahdi Ghandi, Stephen Heil, Prerak Patel, Adam Sapek, Gabriel Weisz, Lisa Woods, Sitaram Lanka, Steven K Reinhardt, Adrian M Caulfield, Eric S Chung, and Doug Burger.
A configurable cloud-scale DNN processor for real-time AI
2018 ACM/IEEE 45th Annual International Symposium on Computer Architecture

FCCM 2017 Andrew G. Schmidt, Gabriel Weisz, and Matthew French.
Evaluating Rapid Application Development with Python for Heterogeneous Processor-based FPGAs
The 25th IEEE International Symposium on Field-Programmable Custom Computing Machines **(Best Short Paper)**

FPGA 2016 Gabriel Weisz, Joseph Melber, Yu Wang, Kermin Fleming, Eriko Nurvitadhi, James C. Hoe.
A Study of Pointer-Chasing Performance on Shared-Memory Processor-FPGA Systems
The 24th ACM International Symposium on Field-Programmable Gate Arrays

FPL 2015 Gabriel Weisz and James C. Hoe.
CoRAM++: Supporting Data-Structure-Specific Memory Interfaces for FPGA Computing
The 25th International Conference on Field Programmable Logic and Applications **(Best Paper)**

FCCM 2014 Eriko Nurvitadhi, Gabriel Weisz, Yu Wang, Skand Hurkat, Marie Nguyen, James C. Hoe, Josè F. Martìnez, and Carlos Guestrin.
GraphGen: An FPGA Framework for Vertex-Centric Graph Computation
The 22nd IEEE International Symposium on Field-Programmable Custom Computing Machines

FPGA 2013	Gabriel Weisz and James C. Hoe.
	C-To-CoRAM: Compiling Perfect Loop Nests to the Portable CoRAM Abstraction
	The 21st ACM International Symposium on Field-Programmable Gate Arrays
FPGA 2012	Eric S. Chung, Michael K. Papamichael, Gabriel Weisz, James C. Hoe, and Ken Mai.
	Prototype and Evaluation of the CoRAM Memory Architecture for FPGA-Based Computing
	The 20th ACM International Symposium on Field-Programmable Gate Arrays

▬▬▬▬ Service

2023	**Workshops and Tutorials Chair** IEEE International Symposium on Field-Programmable Custom Computing Machines (FCCM)
2022	**Publications Chair** IEEE International Symposium on Field-Programmable Custom Computing Machines (FCCM)
2021	**Demo Night Chair** IEEE International Symposium on Field-Programmable Custom Computing Machines (FCCM)
2019-Present	**Technical Program Committee Member** ACM International Symposium on Field-Programmable Gate Arrays (FPGA)
2018-Present	**Technical Program Committee Member** IEEE International Symposium on Field-Programmable Custom Computing Machines (FCCM)
2017-2019	**Technical Program Committee Member** IEEE International Conference on Reconfigurable Computing and FPGAs (ReConFig)
2016-Present	**Technical Program Committee Member** IEEE Conference on Field Programmable Logic and Applications (FPL)
2015	**Grand Awards Judge** Systems Software Category, Intel International Science and Engineering Fair
2013-2015	**Member of the Board of Directors** Cornell Club of Pittsburgh
2012-2015	**Member of the Doctoral Review Committee** Computer Science Department, Carnegie Mellon University
2012-2014	**Member of the PhD Program Admissions Committee** Computer Science Department, Carnegie Mellon University
2012	**Grand Awards Judge** Computer Science Category, Intel International Science and Engineering Fair
2011-Present	**Undergraduate Admissions Contact** Cornell Alumni Admissions Ambassador Network
2004-2005	**Team Mentor** Greater Baltimore Technology Council "Mosh Pit" Business Plan Competition