Fabiano Libano

Bio

Birthday: March 8, 1996 | Birthplace: Madrid, Spain | Nationality: Brazil
Phone: +1 (480) 352-9633
Mail: flibano@asu.edu

Education

Doctor's Degree | Computer Engineering
Arizona State University, ASU, Tempe, United States.
2018 - 2021
Advisor: John Brunhaver
Topic: Low-overhead, algorithm-based hardening techniques and RHBD accelerator architectures for matrix multiplication and neural networks on FPGAs/ASICs.

Master's Degree | Computer Engineering
Universidade Federal do Rio Grande do Sul, UFRGS, Porto Alegre, Brazil
2017 - 2018
Advisor: Paolo Rech
Topic: Reliability analysis and selective hardening of neural networks on FPGAs.

Bachelor's Degree | Computer Engineering
Universidade Federal do Rio Grande do Sul, UFRGS, Porto Alegre, Brazil
2013 - 2017

Languages

Portuguese
Understanding, Speaking, Writing, Reading: Mother Language

English
Understanding, Speaking, Writing, Reading: Fluent

Spanish
Understanding, Speaking, Writing, Reading: Fluent

Main Fields of Interest

1. Fault-Tolerance, Reliability & Radiation Effects
2. Computer Architecture, Digital Systems & FPGAs
3. Artificial Intelligence & Machine Learning
4. Space Exploration
5. Finance

Skills & Expertise

1. RTL & FPGAs: Over 7 years of experience with FPGAs writing and deploying application-specific RTL, tailored for optimal performance (i.e. high throughput, low latency), and efficient power consumption. Particularly, I have a wide familiarity with Xilinx's devices and development tools. Proficient in both Verilog and VHDL.
2. AI & Machine Learning: Experience with modeling and training artificial neural networks for image processing, linear regression, and classification tasks. Additionally, I have experience with reduced precision (i.e. quantized) models and dedicated hardware architectures. I have worked primarily with TensorFlow.
3. Silicon Testing: Well over 1000 hours of hands-on experience with neutron beam testing at RAL-ChipIR (UK) and at LANL-LANSCE (US). I currently work with laser sources for evaluating architectural vulnerabilities factors (AVF), post-silicon. In addition, I also had a brief experience with gamma cells for TID effects at LANL.
4. Languages (from most to least proficient): Python, C, SystemVerilog, Verilog, C++, VHDL, Assembly, Shell, Java, Objective-C, Swift, SQL, JavaScript.
5. Tools (from most to least proficient): Vivado, TensorFlow, Vitis, Avalon, Xcode, Android Studio, SPICE, Virtuoso, Wireshark, MATLAB, Photoshop.

Awards

IEEE NPSS Transactions on Nuclear Science Best Paper

2022

"Selective Hardening for Neural Networks in FPGAs" (#8 on my Publications list) was chosen by TNS editors to receive the 2022 best paper award. Eligible papers were those published in the 3rd year prior (2019), and the winner was the paper judged most useful to the Nuclear Science community according to (a) the rate of downloads and (b) citations in the 3 calendar years including the year of publication.

Publications

1. F. Libano, P. Rech, J. Brunhaver, “Efficient Error Detection for Matrix Multiplication with Systolic Arrays on FPGAs”, at IEEE Transactions on Computers, 2023.
2. F. Libano, P. Rech, B. Neuman, J. Leavitt, M. Wirthlin, J. Brunhaver, “How Reduced Data Precision and Degree of Parallelism Impact the Reliability of Convolutional Neural Networks on FPGAs”, at IEEE Transactions on Nuclear Science, 2021.
3. F. Libano, P. Rech, J. Brunhaver, “On the Reliability of Xilinx’s Deep Processing Unit and Systolic Arrays for Matrix Multiplication”, presented at Radiation and its Effects on Components and Systems (IEEE), 2020.
4. F. Libano, P. Rech, B. Neuman, J. Leavitt, M. Wirthlin, J. Brunhaver, “Evaluating the Impact of Reducing Data Precision on the Reliability of Neural Networks on FPGAs”, presented at Nuclear and Space Radiation Effects Conference (IEEE), 2020.
5. F. Libano, B. Wilson, M. Wirthlin, P. Rech, J. Brunhaver, “Understanding the Impact of Quantization, Accuracy, and Radiation on the Reliability of Convolutional Neural Networks on FPGAs”, at IEEE Transactions on Nuclear Science, 2020.
6. F. Libano, B. Wilson, M. Wirthlin, J. Brunhaver, “Understanding the Impact of Binary Quantization on the Reliability of Neural Networks on FPGAs” presented at Radiation and its Effects on Components and Systems (IEEE), 2019.
7. F. Santos, C. Lunardi, D. Oliveira, F. Libano, P. Rech, “Reliability Evaluation of Mixed-Precision Architectures”, at International Symposium on High-Performance Computer Architecture (IEEE), 2019.
8. F. Libano, B. Wilson, J. Anderson, M. Wirthlin, C. Cazzaniga, C. Frost, P. Rech, “Selective Hardening for Neural Networks in FPGAs”, at IEEE Transactions on Nuclear Science, 2019.
9. F. Benevenuti, F. Libano, V. Pouget, F. Kastensmidt, P. Rech, “Comparative Analysis of Inference Errors in a Neural Network Implemented in SRAM-Based FPGA Induced by Neutron Irradiation and Fault Injection Methods”, presented at Symposium on Integrated Circuits and Systems Design (IEEE), 2018.
10. F. Libano, B. Wilson, M. Wirthlin, P. Rech, “Selective Hardening for Neural Networks in FPGAs”, presented at Nuclear and Space Radiation Effects Conference (IEEE), 2018.
11. F. Libano, P. Rech, L. Tambara, J. Tonfat, F. Kastensmidt, “On the Reliability of Feed-Forward Artificial Neural Networks in FPGAs”, at IEEE Transactions on Nuclear Science, 2017.
12. F. Libano, P. Rech, L. Tambara, J. Tonfat, F. Kastensmidt, “Evaluation of Feed-Forward Artificial Neural Networks Reliability in FPGAs”, presented at Nuclear and Space Radiation Effects Conference (IEEE), 2017.
13. F. Libano, P. Rech, “Reliability Analysis of Feed-Forward Artificial Neural Networks in System on Chips”, presented at Silicon Errors in Logic-Systems Effects (IEEE), 2017.

Work Experience

Intel | Reliability Engineer

2022-Nowadays

I am currently working at Intel's Advanced Reliability Characterization (ARC) group. My NDA limits the extent to which I can share information regarding day-to-day activities and the company's overarching goals from a product reliability standpoint. However, I can state that my position requires extensive architectural comprehension for (a) tracking fault propagation and (b) conceiving mitigation strategies. My work is very experimental in nature: I am in charge of designing and running experiments with post-silicon Engineering Samples. I write sets of highly-automated scripts that interact with real-world equipment (i.e. motors, sensors, lasers) to generate experimental data. I then use common data science libraries (i.e. numpy, pandas, plotly) for parsing and visualization purposes. I also work directly with computer architects on new RAS (reliability, availability, serviceability) features for Xeon.

ASU | PhD Student & Research Assistant

2018-2021

I worked on novel hardening techniques as well as on dedicated RHDB accelerator architectures for matrix multiplication and neural networks on FPGAs/ASICs. Particularly, through extensive experimentation and analysis of multi-level fault models, my research developed error detection/correction methods with minimum added costs (compared to traditional modular redundancy). Moreover, I was involved in designing setups for beam experiments with FPGAs and with a fully custom chip (integrating multiple different compute units and test structures). Tangentially, I was also involved in a project on efficient SW/HW co-design, in which accurate kernel identification (SW) enabled optimal design decisions (HW), for increased performance.

Los Alamos National Laboratory | Intern/Student @ 2019 Radiation Effects Summer School

Summer 2019

The experience at LANL was very unique. I was fortunate to accumulate many hours of hands-on work with different types of radiation experiments (neutron generators for evaluating SEEs, gamma cells for evaluating TID effects, and laser sources for evaluating architectural vulnerabilities). I have learned a lot from a theoretical standpoint as well, since I have attended more than 50 hours of classes and lectures, from world-renowned researchers on particle physics, radiation testing, high-performance computing, space applications and beyond. As my summer project, I have evaluated the impact of reducing floating-point precision of processing elements in neural networks on FPGAs (both in terms of accuracy and radiation sensitivity).

UFRGS | M.S. Student sponsored by CAPES-MEC (Coordination of Improvement of Graduate Level Personnel - Ministry of Education - Brazil)

2017 - 2018

I was specializing as a researcher in reliability and fault tolerance of FPGA and SoC components under harsh conditions, such as radiation-rich environments. The main focus of my dissertation was to evaluate the behavior of neural networks, in safety-critical applications such as autonomous driving and space missions (with a need for real-time image processing). I proposed a novel selective hardening technique for neural networks, which considered the AVF of individual layers for prioritizing modular replication of neurons. The results were presented at international conferences as well as published in IEEE journals.

UFRGS | Intern in the Scientific Initiation Program of CNPq (National Council of Scientific and Technological Development - Brazil)

2015 - 2017

As an undergraduate student, I evaluated the reliability of simple, perceptron-based neural networks for both pattern recognition and linear regression, implemented in resource-constrained FPGAs. The results were presented at IEEE conferences.

Queen Mob | iOS Developer

03/2012 - 10/2012

In parallel with my final year of high school, I briefly interned at QueenMob, where I worked on new iOS apps (such as “Turismo Bento”) and other mobile solutions for regional clients of southern Brazil.

Independent | iOS Developer

2010 - 2013

After the launch of the App Store, I followed my curiosity, studied on my own, and became an iOS developer. Eventually, I launched an app called “iTaxi” which had over 50,000 downloads, coverage by national magazines, newspapers and tv shows.

Open Source Projects

Libano's Systolic Array Generator

https://github.com/lllibano/SystolicArray | https://github.com/lllibano/LABFT

Inspired by Google’s DPU and Xilinx’s TPU, I proposed an INT8 matrix multiplication hardware accelerator for Xilinx FPGAs. Using time-multiplexed DSPs, my architecture outperforms prior work, while utilizing less resources and being more energy efficient. In addition, I wrote a parametric Python script that allows anyone to generate custom-sized arrays that fit their needs and constraints, in a matter of minutes.

References

Ricardo Ascazubi
ricardo.ascazubi@intel.com
Intel Corporation

John Brunhaver
jbrunhaver@asu.edu
Arizona State University

Paolo Rech
prech@inf.ufrgs.br
UFRGS | University of Trento

Heather Quinn
hquinn@lanl.gov
Los Alamos National Laboratory

Michael Wirthlin
wirthlin@byu.edu
Brigham Young University

Christopher Frost
christopher.frost@stfc.ac.uk
Rutherford Appleton Laboratory

Brett Neuman
bneuman@boeing.com
Boeing

Fernanda Kastensmidt
fglima@inf.ufrgs.br
UFRGS