Liana Valdes

Ph.D. Graduate

Knight Foundation School of Computing and Information Sciences

Florida International University (FIU)

Miami, FL, 33199

+1 (786) 665-5542

Email: lvald108@fiu.edu

Web: https://lvald.netlify.app/

GitHub: github.com/lia54

LinkedIn: linkedin.com/in/liana-valdes/

RESEARCH & PROFESSIONAL INTERESTS

Storage Systems, Distributed Systems, Caching Algorithms, ML for Systems, Systems for ML, OS, Team Leadership, Innovation, Motivation, Empathy, Technology, Networking, Creativity, and Relationship Building.

EDUCATION

Florida International University **21 August 2017 - 16 December 2023**

Doctor of Philosophy in Computer Science Miami, FL

Advisor: Eminent Scholar Chaired Professor Raju Rangaswami. GPA: 3.83/4.0

Graduate Relevant Courses: Operating Systems (COP-5614), Analysis of Algorithms (COT-6405), Theory of Computation (COT-5310), Computer Communication and Networking Technologies (TCN-5030), Introduction to Algorithms (COT-5407), Secure Application Programming (CEN-5079), Database Management Systems (COP-5725), Introduction to Machine Learning (CAP-5610), Advanced Software Engineering (CEN-5010), Data Visualization (CAP-5738).

Florida International University **21 August 2017 - 17 December 2022**

Master of Science in Computer Science Miami, FL

Technological University of Havana "José Antonio Echeverría" **1 September 2009 - 20 July 2014**

Bachelor of Science in Telecommunications and Electronics Engineering Havana, Cuba

Faculty of Telecommunications Engineering (FIT). GPA: 4.48/5

Undergraduate Relevant Courses: Linear Algebra and Analytical Geometry, Drafting, Political Economics of Capitalism, Philosophy and Society, Mathematics I, Mathematics II, Molecular Physics and Mechanics, Computerized and Automated Office Systems (Elective), Programming I, Introduction to Engineering Project, Chemistry, Electric Circuits I, Electric Circuits II, Electromagnetism and Optics, Analog Electronics I, Communication Statistics, Quantum Physics, Materials and Components, Probability and Statistics, Advanced Programming (Elective), Programming II, Differential Equations and Series, Sociopolitical Theory, Electrical Circuits III, Digital Electronics I, Digital Electronics II, Analog Electronics II, Analog Electronics III, Principles of Communications I, Principles of Communications II, Virtual Instrumentation (Elective), Transmission Lines, Research Methodology, Numerical Methods, Microprocessors I, Social Issues in Science and Technology, Digital Signal Processing, Field Theory, Antennas, Optical Communications, Telephone Communication, Accounting and Finance, Principal of Communications III, Network Management (Elective), Microprocessors II, Radio Electronics I, Radio Electronics II, Radio Propagation, Networks 1, Networks 2, Network Simulators (Elective), Transmission System, Microprocessor Systems, Business Administration, Satellite Communications (Elective), Principles of Television, Networks 3, and IP Telephone (Elective).

RESEARCH PROJECTS

Extending storage systems to meet cloud needs | Seagate Tech. & SyLab (FIU) **August 2021 - Present**

- Designed and developed TxFuse, a novel architecture to enhance distributed storage systems with features as plugins.
- Designed the transactional coupling and reliable notification mechanisms to develop a taxonomy of plugins efficiently and with less development complexity, and that are deployable as stateless microservices using Docker.
- Evaluated storage plugins, developed on top of the MinIO object store, based on system performance, including latency, throughput, and development complexity, using lines of code (LOC) and McCabe's cyclomatic complexity metric.
- Compared storage features such as Compression, Backup, and Encryption with the same counterparts implemented in the distributed object storage system, MinIO.
- Designed and developed other novel storage features, such as the Semantic enhancer and a Logger.
- Designed the storage-local mechanism to offload plugin functions and run them in the server-side execution context.

A distributed cache for cloud data centers | SyLab, ModLab, DAMRL (FIU) & HASLab (UMinho)   May 2020 - Present

- Designed CaaS, a novel, distributed, shareable, and generalized caching service for multiple types of stores in the cloud.
- Improved hit rate from 4% to 96% compared to a local cache using a write-through policy and running CloudVPS traces.
- Reduced read latency by 5% to 25% and write latency by 99% for applications, services, and microservices that use the caching layer to accelerate performance.
- Designed and implemented a read/write protocol that creates transactions of fractions of writes distributed across the servers to ensure writes are ordered correctly.
- Designed and developed a writeback protocol that writes data back to the storage layer consistently.
- Designed and developed novel QoS algorithms for latency and throughput reservation and optimization.
- Designed and developed a system simulation framework that models a CaaS system, including clients, servers, a coordinator, and the storage layer. Servers simulate the Raft consensus protocol, and the coordinator simulates the Zookeeper consensus protocol.

ML systems for caches in the cloud | SyLab, ModLab, BioRG (FIU) & VISA (ASU)   August 2018 - February 2021

- Analyzed production storage workloads from different sources and cloud service providers.
- Designed and developed ML algorithm LeCaR, a novel cache replacement algorithm that uses reinforcement learning to model the cache replacement problem and achieve outstanding performance when compared with the state-of-the-art.
- Designed and developed a general ML framework, CACHEUS, and its experts, SR-LRU and CR-LFU, that consistently perform best in 47% of cases and have no difference with the best in 40% of cases, while being worse than the compared cache replacement algorithms in 13% of the cases with effect size of less that 0.31.
- Simplified the CACHEUS ML model by eliminating the discount rate and adapting the learning rate hyperparameter.
- Evaluated caching algorithms using an LRU list that uses a combination of a double-linked list and a hash map, and a priority queue using a heap data structure to manage the frequency of items.

PUBLICATIONS

Project Silica: Towards Sustainable Cloud Archival Storage in Glass.   October 2023
Liana Valdes, in Acknowledgement, The 29th ACM Symposium on Operating Systems Principles, SOSP'23.

Infusing Pub-Sub Storage with Transactions   July 2022
Liana V. Rodriguez, John Bent, Tim Shaffer, and Raju Rangaswami, 14th ACM
Workshop, HotStorage'22.

Unifying the Data Center Caching Layer - Feasible? Profitable?   July 2021
Liana V. Rodriguez, Alexis Gonzalez, Pratik Poudel, Raju Rangaswami, and Jason Liu, 13th ACM Workshop,
HotStorage'21.

Learning Cache Replacement with CACHEUS   February 2021
Liana V. Rodriguez, Farzana Yusuf, Steven Lyons, Eysler Paz, Raju Rangaswami, Jason Liu, Ming Zhao, and Giri Narasimhan, 19th
USENIX Conference, FAST'21.

Driving Cache Replacement with ML-Based LeCaR   July 2018

Giuseppe Vietri, Liana V. Rodriguez, Wendy A. Martinez, Steven Lyons, Jason Liu, Raju Rangaswami, Ming Zhao, Giri Narasimhan.
USENIX Workshop, HotStorage'18.

PRESENTATIONS

CORTX and FDMI | CORTX Meet the Architect Series   November 2022
Liana V. Rodriguez and John Bent.

Infusing Pub-Sub Storage with Transactions | HotStorage'22 Presentation   July 2022
Liana V. Rodriguez, John Bent, Tim Shaffer, and Raju Rangaswami.

Unifying the Data Center Caching Layer - Feasible? Profitable? | HotStorage'21 Presentation   July 2021
Liana V. Rodriguez, Alexis Gonzalez, Pratik Paudel, Raju Rangaswami, and Jason Liu.

Learning Cache Replacement with CACHEUS | Poster Annual FIU SCIS Research Day   October 2019
Liana V. Rodriguez, Farzana Yusuf, Steven Lyons, Eysler Paz, Raju Rangaswami, Jason Liu, Ming Zhao, and Giri Narasimhan.

Driving Cache Replacement with ML-Based LeCaR | HotStorage'18 Poster session   July 2018
Giuseppe Vietri, Liana V. Rodriguez, Wendy A. Martinez, Steven Lyons, Jason Liu, Raju Rangaswami, Ming Zhao, and Giri Narasimhan.

Graduate Research Assistant/ Teaching Assistant/ GAANN Fellowship | FIU, FL, US August 2017 - December 2023

- Analyzed production storage workloads and discovered workload patterns called workload primitives. I studied how these workloads affect the cache in the real world using a classification based on the composition of primitive workload types. I computed metrics to characterize the production storage workloads. For instance, for a one-day trace from MSR, Cambridge, 3.29% of the requests were write operations, and the unique and reuse footprints were 89.24% and 2.22%, respectively. The research dives into the intricacies of the data and examines the effects of several factors, providing valuable insights. This approach has improved cache performance in novel cache replacement algorithms capable of adapting to the data and tailored solutions to meet the needs of the production storage workload.

- Designed ML systems for cache replacement, LeCaR, and CACHEUS, which enhance storage caching performance. When evaluating a sample trace from MSR, Cambridge, LeCaR, and CACHEUS algorithms outperform state-of-the-art cache replacement algorithms, including LRU and ARC, with hit-rate values of 93.61% and 94.38%, respectively.

- Designed CaaS, a distributed caching service that improves hit rate, service, and microservice reading and writing performance compared with multi-tenant caches and local caches. CaaS outperforms an SSD-based local cache using a write-through policy. We addressed data consistency, scalability, high availability, load balancing, disaster recovery, and fault tolerance to guarantee optimal performance and reliability.

- Designed and developed TxFuse, a novel architecture to develop novel storage features as plugins. Object storage systems such as MinIO are highly complex. For instance, MinIO includes 397 releases with six storage features: Erasure coding, Encryption, Disk cache, Compression, Lambda transformation, and Replication implemented as an integrated part of the storage core codebase. TxFuse enables the development and deployment of storage system features independent of the storage system using two mechanisms: transactional coupling and reliable notification. It additionally allows the offloading of plugin functions to reduce the network overhead and client operation latency through the storage-local mechanism that decouples the physical resources of the features from the storage system. When compared with native MinIO counterparts, TxFuse reduces code size by 2.82X for Compression, 27.91X for Backup, and 26.54X for Encryption.

- Taught Operating Systems course at the undergraduate level in Fall 2018 and Spring 2019. Activities include grading projects, office hours for students, and assisting students in the Nachos instructional OS.

Research Intern | Seagate Technology, Remote, US   August 2021 - December 2021

- Seagate, a leading data storage company, developed CORTX, an open-source object storage solution. The solution facilitates the storage and management of massive amounts of unstructured data, such as media files, logs, and backups. CORTX has an integrated software stack comprising several components. I tested the CORTX software stack to ensure seamless integration and deployment.

- Tested the CORTX's interface, File Data Manipulation Interface (FDMI), which uses plugins to implement storage features. I developed plugins such as a word count program that counts the number of times a word appears in each S3 PUT request to the storage. This plugin is equivalent to a program that polls a directory every 3 seconds and prints the 30 most popular words within each file. This research contributes to the design and development of more robust and efficient storage systems that can meet the ever-evolving needs of the modern world.

- Conducted a performance study on Motr, a distributed object storage system with advanced features and capabilities deployed at the Jülich Supercomputing Center (JSC) in Germany. The study aimed to analyze the system's performance and evaluate its efficiency. We reported the maximum achievable network and device bandwidth in a Sage platform x86 Motr cluster with sixteen clients and seven server nodes; each server had three tiers (tier-1, tier-2, and tier-3) that had NVMe, SSD, and HDD storage devices, respectively. The results showed 99% and 86% bandwidth utilization for read and write operations in tier-2 compared to the theoretical system limit.

Research Intern | Microsoft Research (MSR), Cambridgeshire, UK   January 2020 - March 2020

- Helped develop the software stack for a storage system using cutting-edge quartz glass technology as part of Microsoft's Project Silica team. This innovative technology is specifically designed to cater to the requirements of Azure Cloud archival workloads, enabling efficient and cost-effective storage. Unlike traditional storage systems, Project Silica's approach involves storing vast amounts of cold data, creating tiny 'voxels' in quartz glass.

- Designed ML systems using Deep Learning and Unsupervised Learning that can identify anomalies when tested using data from the Warner Bros. movie "Superman," written on glass. These ML models can identify unusual patterns in the image data, particularly fine-grained errors within one sector (XY plane) and coarse-grained errors within one track (Z direction).

Telematics Specialist B | Cuban Telecommunications Company S.A.   September 2015 - June 2016

- Maintained uninterrupted and efficient communication services, crucial for monitoring the core infrastructure of PSTN, ATM, DSL, and public and private VoIP telephone systems. This involves closely observing the hardware, software, and network elements that comprise these systems' backbone. Any potential issues were detected and resolved promptly by continuously monitoring performance metrics such as speed, latency, uptime, and number of lines interrupted.

- Monitored the security of these systems to prevent unauthorized access and data breaches. With a focus on proactive monitoring and maintenance, these critical communication systems were highly reliable and available, ensuring seamless communication for businesses and individuals.

- Helped conduct a network assessment to ensure the continuous deployment of the Fiber to the Home (FTTH) networks in Old Havana. The assessment helped to identify the best action plan for deploying OTN devices.

- Helped assess and enhance PSTN network infrastructure to meet the community's needs.

Research Intern/Apprenticeship | Cuban Radio and Television Broadcasting Company   March 2012 - July 2014

- Analyzed the board layout to modify the UHF and VHF communication modules from PAL to NTSC television standards. This modification improved the frequency of the television signal, resulting in a much more precise and higher-quality picture.

- Helped to design RF matching networks that use passive components to achieve maximum power transfer to a 50-ohm or a real-world load, such as a television antenna. This involves analyzing the properties of the RF signal and determining the best arrangement of passive components to match the load's impedance.

- Simulated results obtained with MATLAB, virtual instrumentation of E-field sensors, and spectrum/network analyzers. The simulations allowed us to visualize the RF signal's properties, including frequency, amplitude, and phase.

HONORS & AWARDS

USENIX Student Travel Award, FAST'18 & FAST'19 & FAST'23.	**Sep. 2018, Sep. 2019, and Sept. 2023**
GAANN Fellowship from the U.S. Department of Education (ED).	**Jan. 2022 & Jan. 2023**
CMD-IT/ACM Richard Tapia Celebration of Diversity in Computing Conference Scholarship.	**Sep. 2022 & Sep. 2024**
Grace Hooper Celebration of Women in Computing FIU Scholarship.	**Oct. 2019 & Sep. 2022**
USENIX SREcon24 Europe/Middle East/Africa Conference Award	**Nov. 2024**

LEADERSHIP & ACTIVITIES

- Selected as SyLab team leader and team leader for three research projects at FIU.

- Selected to plan and organize an event for Seagate's CORTX Global Hackathon, focusing on integration challenges, potential solutions, and open-source software.

- Selected as a national anonymous reviewer for peer-reviewed journal articles for ACM Transactions on Storage (TOS), the leading journal for publishing advances in data storage research, in 2021 and 2025.

- Selected presentation slides for an annual Women in Computer Science (WiCS) meeting, focused on sharing experiences and strategies for achieving academic success in computer science.

- Participated in Upsilon Pi Epsilon (UPE) activities, including workshops, presentations, tutorials, and leadership events to develop leadership skills, build industry connections, and achieve academic success.

- Selected as a volunteer for an eight-month teaching program for Google's CS First program, focusing on coding skills and learning computer science concepts using Scratch for children under eight years old.

TEACHING ASSISTANT

Operating Systems	**August 2018 & January 2019**
Florida International University	Miami, FL
Transmission Systems II	**2011 - 2012**
Technological University of Havana "José Antonio Echeverría"	Havana, Cuba
Analog Electronics I	**2010 - 2011**
Technological University of Havana "José Antonio Echeverría"	Havana, Cuba

Technical Skills

Skills: Caching, Storage Systems, Machine Learning, Reinforcement Learning, Deep Learning, Supervised Learning, Unsupervised Learning, Artificial Intelligence, Distributed Systems, Distributed caching, QoS algorithms.
Languages: Python, C#, C, C++, Java, Go, R, R++, TeX, Visual Basic, High-Level Assembly (HLA).

Tools/Frameworks: Git, GitLab, Docker, VirtualBox, VMware Workstation, VS Code, Tableau, MATLAB, LabVIEW, LaTeX, PyTorch, TensorFlow, Scikit-learn.

Leadership Skills

Team Leadership, Relationship Building, Emotional Intelligence, Forward-thinking, Problem-solving, Empowerment, Adaptability, Motivation, Innovation, Inspiration, Active Listening, Empathy, Presentation, Authenticity, Honesty, Creativity, Integrity, Decision-making, Accountability, Strategic Thinking, Communication, Coaching, and Management.

Certifications

CompTIA A+ 220-801 Certification Exam.

CompTIA A+ 220-802 Certification Exam.

Language

Spanish: Fluent.

English: Full Professional Proficiency.

Hobbies

Music: Traditional Cuban Music, Country Music, Pop Music, Blues, and Jazz.
Poetry: Cuban Poetry, Latino-American Poetry, Spanish Poetry, and American Poetry.
Chess: Tournaments, Titles and Rankings.
Fitness: Physical Fitness, Exercise, Training, and Walking.
Animation: Computer Animation, Disney Animation, and Japanese Animation.