

Liana Valdes Rodriguez

Ph.D. Graduate
Knight Foundation School of Computing and Information Sciences
Florida International University
Miami, FL, 33199

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

SUMMARY

As a Ph.D. graduate student in Computer Science (CS) at the Systems Research Laboratory (SyLab) under the supervision of **Professor Raju Rangaswami**, I have been exploring the fascinating realms of storage systems, distributed systems, operating systems (OS) and the intersection between systems and ML. In my previous research projects, I successfully led a team focused on designing and implementing novel caching algorithms for storage caches, significantly improving performance. I have also actively developed and enhanced distributed systems solutions for today's cloud data centers (DC).

Before graduating from my Ph.D. program at **Florida International University (FIU)**, I gained valuable industry experience. I was a research intern at Microsoft Research (MSR) Cambridge, UK, where I contributed to the Project Silica. In 2021, I further honed my skills as an intern at Seagate Technology. These experiences have enriched my understanding of the field and provided practical insights to apply to my current research. Before starting my Ph.D., I completed the BSc in Electronics and Telecommunications Engineering in Havana, Cuba.

EDUCATION

Florida International University (FIU)

21 August 2017 - 16 December 2023

Doctor of Philosophy in Computer Science (CS)

Miami, FL

Advisor: Eminent Scholar Chaired Professor Raju Rangaswami.

GPA: 3.83/4.0

Graduate Relevant Courses: Operating Systems (OS, 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), Data Base Management Systems (DBMS, COP-5725), Introduction to Machine Learning (ML, CAP-5610), Advanced Software Engineering (CEN-5010), Data Visualization (CAP-5738).

Florida International University (FIU)

21 August 2017 - 17 December 2022

Master of Science in Computer Science (CS)

Miami, FL

Advisor: Eminent Scholar Chaired Professor Raju Rangaswami.

Technological University of Havana “José Antonio Echeverría”, CUJAE 1 September 2009 - 28 July 2014

Bachelor of Science in Electronics and Telecommunications 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, Network 3, and IP Telephone (Elective).

RESEARCH PROJECTS

- Extensible Distributed Storage Systems | Seagate Technology & SyLab (FIU) August 2021**
- Developed TxFuse architecture that uses feature plugins to enhance a distributed storage system.
 - Created a plugin taxonomy based on transactional coupling and reliable notification.
 - Evaluated plugin prototypes based on performance and development complexity.
- Distributed Caches in Data Centers | SyLab, ModLab, DAMRL (FIU) & HASLab (UMinho) May 2020**
- Developed CaaS, a distributed caching service for cloud storage production workloads.
 - Improved the latency and performance of read and write operations compared to back-end storage.
 - Developed a simulation framework to model the CaaS system's various components.
- Caching Algorithms for Storage Caches | SyLab, BioRG (FIU) & VISA (ASU) August 2017**
- Analyzed production storage workloads from different cloud service providers.
 - Developed novel cache replacement algorithms using ML techniques to improve cache hit rate for cloud workloads.
 - Reduced the cost of accessing SSD-based caches by reducing the total of writes sent to the devices.

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.
- To Cache or Not to Cache**
Steven Lyons, Liana V. Rodriguez, and Raju Rangaswami.
- 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, and 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.

HONORS & AWARDS

USENIX Student Travel Award, FAST'18 & FAST'19 & FAST'23.

September 2018, 2019, and 2023

GAANN Fellowship from the U.S. Department of Education.

January 2022 & January 2023CMD-IT/ACM Richard Tapia Celebration of Diversity in Computing Conference Scholarship. **September 2022**Grace Hooper Celebration of Women in Computing FIU Scholarship. **October 2019 & September 2022**

EXPERIENCE

Team Leader | Systems Research Laboratory (SyLab), FL, US**August 2018 - December 2023**

- I worked on setting and communicating the teams' goals effectively, developing and implementing plans and strategies, and delegating responsibilities to team members. Additionally, I coached members on achieving goals and developing the necessary skills and preparation to get the most effective results for the project. I also worked when roadblocks inevitably emerged while resolving a project problem, analyzing problems objectively, brainstorming solutions, evaluating alternatives, and implementing the optimal approach.
- I regularly monitored the team's performance, measuring their output, reviewing their progress against set targets, and providing feedback where necessary. I identified the inefficiencies and areas that required improvement by evaluating how the team completed tasks and how well they achieved their goals.
- As a team leader, I helped motivate team members and worked to keep them engaged whenever I noticed a decline in their enthusiasm so that the team's productivity grew.

Graduate Research Assistant / GAANN Fellow | SyLab, FL, US**August 2017 - December 2023**

- I analyzed production storage workloads and discovered workload patterns called workload primitives. As a result, I analyzed how these workloads function in the real world using a classification based on the composition of primitive workload types. Additionally, I computed metrics to characterize the production storage workloads. For example, for a one-week trace from MSR Cambridge, 3.29% of the requests were writes, the unique footprint was 89.24%, and the reuse footprint was 2.22%. As a result, my research dives into the intricacies of the data and examines the effects of different factors, providing valuable insights for professionals in the field. This approach has improved cache performance in novel cache replacement algorithms capable of adapting to the data and demonstrated our ability to tailor solutions to meet workload needs.
- Furthermore, we utilized ML techniques to develop two novel cache replacement algorithms, LeCaR and CACHEUS, that enhance storage caching performance. When evaluating a one-week trace from MSR Cambridge, these algorithms outperform state-of-the-art caches with hit-rate values of 91.25% and 92.5%, respectively, for LeCaR and CACHEUS.
- We developed a distributed caching service called CaaS that outperforms other caches in the cloud data center, caches reads, and uses writeback/writes-around policies for writes. CaaS outperforms an SSD-based local cache, with a hit rate of 96% for a one-day trace from Cloud VPS. In these systems, ensuring data consistency, scalability, high availability, failover, disaster recovery, and fault tolerance must be addressed to guarantee optimal performance and reliability.

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

- Seagate, a leading data storage company, has developed CORTX, an open-source object storage solution. The solution facilitates storing and managing massive amounts of unstructured data, such as media files, logs, and backups. CORTX has an integrated software stack comprising several components that work harmoniously to provide a complete object storage solution. I worked on conducting testing to ensure seamless integration and deployment of the software stack. This includes deploying cortx-motr, cortx-hare, cortx-ha, and cortx-s3server. I also worked on organizing the Seagate Global Hackathon and suggesting possible solutions to the challenges.
- I worked on the CORTX Storage Extensible Interface called File Data Manipulation Interface (FDMI), which uses plugins that implement storage features. I developed examples of plugins, such as a word count program that counts the number of times a word appears in every 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 developing more robust and cutting-edge storage that can meet the ever-evolving needs of the modern world.

- I conducted a performance study at the Jülich Supercomputing Center (JSC) in Germany on Motr, a distributed object storage system with advanced features and capabilities. 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 16 client nodes and 7 server nodes; each server had three tiers (tier-1, tier-2, and tier-3) that have 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), Cambridge, UK

January 2020 - March 2020

- As a member of the Project Silica team at Microsoft, I had the opportunity to work on the Cloud Systems Futures cloud initiative. Specifically, my role involved working on a software stack for a storage system that utilized cutting-edge quartz glass technology. 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 creating tiny 'voxels' into quartz glass using a femtosecond laser, which can contain vast amounts of data.
- To improve the storage system's software pipeline's error analysis capabilities in Project Silica, I developed ML techniques using Python. Specifically, I used isolation forests and encoders as anomaly detection algorithms that can help identify unusual patterns in the data, particularly fine-grained errors within one sector (XY plane) and coarse-grained errors within one track (Z direction).
- I developed ML models that can analyze large amounts of data and identify anomalies when tested using real data from the Warner Bros. movie "Superman" written to the glass.

Network Engineer | Telecommunications Company (ETECSA)

September 2015 - June 2016

- I 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.
- It was also essential to monitor 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.
- We conducted a network assessment to ensure the continuous deployment of the fiber-to-the-home (FTTH) networks in Old Havana. This assessment allowed us to identify the best course of action for deploying the Optical Transport Network (OTN) devices. By utilizing Gigabit Passive Optical Networking (GPON) and Ethernet Passive Optical Networking (EPON) variations of Passive Optical Networking (PON) technology, we were able to establish and enhance network infrastructure that meets the community's needs.

Intern | Radiocommunications Company (RadioCuba)

March 2012 - July 2014

- We made significant changes to the board layout to modify the Ultra High Frequency (UHF) and Very High Frequency (VHF) communication modules from the Phase Alternate Line (PAL) to the National Television Standard Committee (NTSC) television standards. This modification improved the frequency of the television signal, resulting in a much more precise and higher-quality picture.
- I designed radiofrequency (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 (TV) antenna. This involves analyzing the properties of the radiofrequency (RF) signal and determining the best arrangement of passive components to match the load's impedance.
- I used MATLAB simulations, which modeled the results, and used virtual instrumentation of E-field sensors and spectrum/network analyzers. The simulations allowed us to visualize the radiofrequency (RF) signal's properties, including frequency, amplitude, and phase.

LEADERSHIP & ACTIVITIES

- Selected to plan and organize a one-day event for Seagate's global CORTX Hackathon, focusing on integration challenges, possible solutions, and open-source object storage software.
- Selected as a participant and anonymous reviewer nationwide for a peer-reviewed journal article for the ACM Transactions on Storage (TOS), the premier journal for publishing advancements in data storage research.
- Selected as a presenter for an annual Women in Computer Science (WiCS) meeting, focusing 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 8-month teaching program for the Google CS First initiative, focusing on coding skills and learning computer science concepts using the programming language called Scratch for kids under 8.
- Selected as the team leader for four research projects at the Systems Research Laboratory at Florida International University.

TEACHING ASSISTANT

Operating Systems (Undergraduate course)

Florida International University

August 2018 & January 2019

Miami, FL

Transmission Systems II

Technological University of Havana “José Antonio Echeverría”, CUJAE

2011 - 2012

Havana, Cuba

Analog Electronics I

Technological University of Havana “José Antonio Echeverría”, CUJAE

2010 - 2011

Havana, Cuba

TECHNICAL SKILLS

Languages: Python, C#, C, C++, Go, R, R++, TeX, Visual Basic, High-Level Assembly (HLA).

Tools/Frameworks: GitLab, Docker, VirtualBox, VMware Workstation, VsCode, Tableau, MATLAB, LabVIEW, LaTeX, Pytorch, Scikit-learn.

CERTIFICATIONS

CompTIA A+ 220-801 Certification Exam.

CompTIA A+ 220-802 Certification Exam.

LANGUAGE

Spanish: Native.

English: Fluent and Full Professional Proficiency.

HOBBIES

Music: Traditional Cuban Music, Country Music, Pop Music, Blues, and Jazz.

Poetry: Cuban Poetry, Latino-American Poetry, Spanish Poetry, American Poetry.

Chess: Tournaments and Matches, and Titles and Rankings.

Fitness: Physical Fitness, Exercise, Training, Walking.

Animation: Computer Animation, Disney Animation, Japanese Animation.

LEADERSHIP SKILLS

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.