Liana Valdes Rodriguez

Knight Foundation School of Computing and Information Science
Florida International University
+1 (786) 665-5542 | lvald108@fiu.edu | linkedin.com/in/liana-valdes | github.com/lia54

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 newly graduated Ph.D. student in **Computer Science (CS)** at the **Systems Research Laboratory (Sylab)** under the supervision of **Prof. Raju Rangaswami**, I have been exploring the fascinating realms of storage systems, distributed systems, operating systems, 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, leading to a significant improvement in performance. I have also been actively involved in the development and enhancement of distributed systems solutions for today's cloud **data centers (DC)**. Prior to my Ph.D. program at **Florida International University (FIU)**, I had the privilege of gaining 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 a research intern at **Seagate Technology**. These experiences have not only enriched my understanding of the field, but have also provided me with practical insights that I can apply to my current research. Before embarking on my Ph.D. journey, 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 (CUJAE)　　　　　　　**1 September 2009 - 28 July 2015**
Bachelor of Science (BSc) 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).

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   Jul. 2022

Liana V. Rodriguez, John Bent, Tim Shaffer, and Raju Rangaswami.

Unifying the Data Center Caching Layer - Feasible? Profitable? | HotStorage'21 Presentation Jul. 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.

ANX: Caching with Anxiety | Poster Annual FIU SCIS Research Day   October 2019

Steven Lyons, Liana V. Rodriguez, and Raju Rangaswami.

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.

EXPERIENCE

Graduate Research Assistant (GRA) | SyLab, Florida, US   **August 2017 - April 2023**

- I have thoroughly analyzed actual production storage workloads and discovered some fascinating insights. By utilizing advanced data analysis methods, I have comprehensively understood how these workloads function in the real world. My research dives into the intricacies of data transfer rates and examines the effects of different types of file compression, providing valuable insights for professionals in the field. This approach has improved performance and demonstrates our ability to tailor solutions to meet workload needs. Besides, we have utilized ML techniques to develop novel caching algorithms that enhance storage caching performance.

- We possess significant expertise in working with distributed systems, which are computer systems of numerous interconnected components that communicate and coordinate their actions via a network. In these systems, ensuring data consistency, scalability, high availability, failover, disaster recovery, and fault tolerance are critical challenges that must be addressed to guarantee optimal performance and reliability. Our experience in distributed systems allows us to develop and implement solutions that tackle these challenges, thereby ensuring that our clients' systems are dependable, efficient, and secure.

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 storage systems 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 data storage. Unlike traditional storage systems, Project Silica's approach involves etching tiny 'voxels' into quartz glass using a femtosecond laser, which can contain vast amounts of data. Through the co-design of hardware and software, this system offers enhanced vulnerability management, ensuring your data remains protected for years to come.

- To improve the data recovery and error analysis capabilities of a storage system's software pipeline, it is recommended to develop ML techniques using Python, PyTorch, Scikit-learn, Isolation Forest, and Encoders. PyTorch is a deep learning framework that simplifies the development of complex neural networks. Scikit-learn is a ML library that provides tools for classification, regression, and clustering tasks. Using these tools in combination, you can develop powerful ML models that can analyze large amounts of data and identify patterns that may be missed by traditional methods. By implementing these models into the storage system's software pipeline, you can improve the accuracy and speed of data recovery and error analysis, leading to a more efficient and reliable system.

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 is designed to facilitate 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 that work harmoniously to provide a complete object storage solution. I worked on conducting thorough testing to ensure seamless integration and deployment of the software stack. Seagate is committed to delivering a reliable and scalable solution that meets the needs of its customers. I also had the pleasure of organizing the Seagate Global Hackathon. I researched and suggested possible solutions to the challenges, which could help the participants develop their ideas and turn them into viable projects. The event was a resounding success, with participants showcasing their creativity and expertise.

- I thoroughly analyzed the CORTX Storage Extensible Interface, called the File Data Manipulation Interface (FDMI). During my research, I utilized C and Python programming languages to implement distributed consensus and transactions. By adopting this approach, we can enhance the scalability and fault tolerance of the storage system, making it more resilient and reliable. This research aims to contribute to developing more robust and cutting-edge storage solutions that can meet the ever-evolving needs of the modern world. Besides, I conducted a comprehensive performance study at the Jülich Supercomputing Center (JSC) in Germany on Motr, an object storage system with advanced features and capabilities. Through distributed systems, we can ensure the integrity and availability of data, even in the face of challenging conditions. The study aimed to analyze the system's performance in-depth and evaluate its efficiency in handling large data sets.

Network Engineer | Telecommunications Company (ETECSA)   September 2015 - June 2016

- I have experience in ensuring uninterrupted and efficient communication services, which are crucial for monitoring the core infrastructure of ATM, DSL, PSTN, 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 can be detected and resolved promptly by continuously monitoring performance metrics such as speed, latency, and uptime. It's 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 can maintain high reliability and availability, ensuring seamless communication for businesses and individuals.

- To ensure the success of the Fiber to the Home (FTTH) networks in Old Havana, we conducted a thorough network assessment. This assessment allowed us to identify the best course of action for deploying the necessary 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 needs of the community.

Intern | Radiocommunications Company (RadioCuba)   March 2012 - July 2015

- 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 was made to improve the frequency of the television signal, and the results were quite impressive. With the new board layout, the frequency of the television signal has been dramatically enhanced, resulting in a much more precise and higher-quality picture.

- I have experience designing 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 RF signal and determining the best arrangement of passive components to match

the load's impedance. To ensure the accuracy of the design, I used MATLAB simulations, which modeled the results using E-field sensors and spectrum/network analyzers. The simulations allowed me to visualize and analyze the RF signal's properties, including frequency response, amplitude, and phase.

RESEARCH PROJECTS

Caching Algorithms for Storage Caches | SyLab, BioRG (FIU) & VISA (ASU)   August 2017
- Analyze storage workloads of different cloud service providers.
- Develop new cache replacement algorithms to improve the hit rate for cloud application workloads.
- Reduce the cost of accessing SSD-based caches by reducing the total of writes sent to the devices.

Distributed Caches in Data Centers | SyLab, ModLab, DAMRL (FIU) & HASLab (UMinho)   May 2020
- Develop CaaS, a distributed caching service for cloud storage production workloads.
- Improve the latency and performance of cache read and write operations compared to the back-end storage.
- Develop a simulation framework to model a writable caching system's various components.

Extensible Distributed Storage Systems | Seagate Technology & SyLab (FIU)   August 2021
- Develop TxFuse architecture to enhance CORTX's distributed storage system with plugins.
- Create a plugin taxonomy that utilizes transactional coupling and a reliable notification system.
- Assess plugin prototypes based on performance and development complexity.

AWARDS

CMD-IT/ACM Richard Tapia Celebration of Diversity in Computing Conference Scholarship in **2022.**
Grace Hooper Celebration of Women in Computing FIU Scholarship in **GHC'19** & **GHC'22.**
GAANN Fellowship from U.S. Department of Education in **2022** and **2023.**
USENIX Student Travel Award in **FAST'23** & **FAST'19** & **FAST'18.**

TEACHING ASSISTANT

Operating Systems (Graduate)   August 2018 & January 2019
Florida International University   Miami, FL

Transmission Systems II   2011 - 2012
Technological University of Havana "José Antonio Echeverría", CUJAE   Havana, Cuba

Analog Electronics I   2010 - 2011
Technological University of Havana "José Antonio Echeverría", CUJAE   Havana, Cuba

SOCIETIES & ACTIVITIES

I was a planner, presenter, and team leader at Florida International University (FIU) from **2018** to **2023**.
Planner of the Seagate Global CORTX Hackathon Event in **2021**.
Reviewer for the ACM Transactions on Storage (TOS) Journal in **2021**.
Presenter for the Women in Computer Science (WiCS) meeting in **2019**.
Participated in UPsilon Pi Epsilon (UPE) activities on campus in **2019, 2021,** and **2022**.
Volunteered for the Google Computer Science (CS) First program in **2019** as a coding instructor for kids under 8.

TECHNICAL SKILLS

Languages: Python, C#, C, C++, Go, R, R++, TeX, Visual Basic, High Level Assembly (HLA).
Tools: GitLab, Docker, VirtualBox, VMware Workstation, VsCode, Tableau, MATLAB, LabVIEW, LaTeX.

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, Latinoamerican Poetry, Spanish Poetry, American Poetry, and Russian Poetry.
Chess: Chess Play Online Application with Time Control, Tournaments and Matches, and Titles and Rankings.
Fitness: Physical Fitness, Exercise, Training, Walking.
Animation: Computer Animation, Disney Animation, Japonese Animation, and Russian Animation.

Skills

Relationship Building, Emotional Intelligence, Forward-thinking, Empowerment, Adaptability, Motivation, Innovation, Inspiration, Active Listening, Empathy, Presentation, Authenticity, Honesty, Creativity, and Management.