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

Research Interests

Storage, distributed systems, caching algorithms, ML for systems, systems for ML, and OS.

Education

Florida International University **21 August 2017 - 17 December 2022**
Master of Science in Computer Science *Miami, FL*
Advisor: Eminent Scholar Chaired Professor Raju Rangaswami

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, Analysis of Algorithms, Theory of Computation, Computer Communication and Networking Technologies, Introduction to Algorithms, Secure Application Programming, Principles of DBMS, Introduction to Machine Learning, Advanced Software Engineering, Data Visualization.

Technological University of Havana **1 September 2009 - 28 July 2015**
Bachelor of Science in Electronic and Telecommunications Engineering *Havana, Cuba*
Faculty of Telecommunications Engineering (FIT) GPA: 4.48/5

Undergraduate Relevant Courses: Linear Algebra and Analytical Geometry, Mathematics I, Mathematics II, Molecular Physics and Mechanics, Computerized and Automated Office Systems, Programming I, Chemistry, Electric Circuits I, Electric Circuits II, Electromagnetism and Optics, Analog Electronics I, Communication Statistics, Probability and Statistics, Advanced Programming, Programming II, Differential Equations and Series. Electrical Circuits III, Digital Electronics I, Digital Electronics II, Analog Electronics I, Analog Electronics II, Principles of Communications I, Principles of Communications II, Virtual Instrumentation, Transmission Lines, Numerical Methods, Microprocessors I, Digital Signal Processing, Field Theory.

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, and Giri Narasimhan, USENIX Workshop, HotStorage'18.

Presentations

CORTX and FDMI | CORTX Meet the Architect Series **November 2022**

Liana V. Rodriguez and John Bent

Learning Cache Replacement with Cacheus | Poster First 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 First 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 | SyLab, Florida, US **August 2017 - April 2023**

Florida International University

- Developed novel caching algorithms for storage caches using ML techniques to improve performance.
- Analyzed real-world production storage workloads to identify application I/O patterns that can be exploited in the design of novel caching algorithms customized to certain applications which improves performance.
- Research in distributed systems that addresses distributed storage challenges such as data consistency, high scalability, high availability, failover, disaster recovery, and fault tolerance.

Research Intern | Microsoft Research, Cambridge, UK **January 2020 - March 2020**

- I worked on Microsoft's Project Silica as part of the optics for the cloud initiative called Cloud Systems Futures developing a storage software stack for storage systems based on quartz glass technology.
- Develop ML techniques in the software pipeline that work to improve data recovery and error analysis in the storage system (**Python, PyTorch, Scikit-learn, Isolation Forest, Encoders**).
- Project Silica aims to develop the first Azure Archival Storage based on solid-silica glass and laser-induced femtosecond pulse technology for Azure Cloud archival workloads.
- The archival storage system developed in Project Silica co-designs hardware and software with a focus on sustainability and vulnerability.

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

- Seagate's object storage solution, CORTX testing integration, and deployment of the software stack.
- CORTX Storage Extensible Interface research (**C, Python, Distributed Consensus & Transactions**).
- I worked on planning the Seagate Global Hackathon event in 2021 where I defined and created the challenge requirement, steps to follow, and possible solutions.
- Performance study of Motr, an object storage system deployed at the Jülich Supercomputing Center (JSC) in Germany (**C, Go, fio, SelfNet Lustre**).

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

- Monitor the core infrastructure of the ATM, DSL, and PSTN public and VoIP telephone systems for a high telephone network reliability and availability.
- Performed network assessment to deploy OTN devices for the EPON/GPON fiber-to-home project in Old Havana.

Intern | Cuba's Radiocommunication Company (RadioCuba) **March 2012 - July 2015**

- Modify UHF/VHF modules from PAL to NTSC television standards by changing the board layout to improve the frequency of the television signal.
- Designed RF matching networks with passive components for maximum power transfer to 50-ohm load or real-world load (TV antenna) and model results with MATLAB simulations using E-field sensors and spectrum/network analyzers. (**MATLAB, Network Analyzers**).

Research Projects

Caching Algorithms for Storage Caches | SyLab, BioRG (FIU) & VISA (ASU) August 2017

- Characterize production storage workloads from different cloud service providers.
- Design and develop new cache replacement algorithms to improve the hit rate performance of caches that store data for cloud application workloads. We developed multiple algorithms, in particular, two algorithms: LeCaR and CACHEUS outperform classical cache replacement algorithms. CACHEUS is inspired by LeCaR but overcomes an important flaw by being completely adaptive, with the elimination of all statically chosen hyperparameters, thus guaranteeing high flexibility.
- Identify cache-relevant characteristics that are reported by workload primitive types. We identify four primitive types of workload: LRU-friendly, LFU-friendly, scan, and churn. Workload primitive types vary between workloads, within a single workload over time, and as cache size changes.
- Reduce the cost of accessing SSD-based caches by reducing the total of application writes sent to storage devices (SATA SSD and NVMe SSD).

Distributed Caches in Data Centers | SyLab, ModLab, DAMRL (FIU) & HASLab (UMinho) May 2020

- Design and develop Caching-as-a-Service (CaaS), a distributed and generalized caching service that addresses the requirements of different types of cloud storage production workloads. First and foremost, CaaS is designed as a general cache utility for a variety of store types, including block storage, object storage, file systems, and key-value stores. CaaS integrates into the storage I/O software stack in cloud infrastructure services and is entirely transparent to applications.
- Design and develop CaaS as a writable cache. To achieve durability, fault tolerance and scalability, CaaS stores data in a set of nodes, each implemented as a fault-tolerant cluster managed by Raft, a distributed consensus protocol.
- Improve cache read and write latency and performance relative to back-end storage. CaaS performs write-back calls eagerly to improve performance. It also proactively schedules asynchronous eviction calls to deliver faster cache access time to the application and increase available cache space which is then used to cache new writes.
- Implement a simulation framework to simulate the different components of a consistent and writable caching system. The CaaS simulator is a full system simulator that implements the operation of multiple types of components, including clients, servers, the coordinator, and network communication between them.

Extensible Distributed Storage Systems | Seagate Technology & SyLab (FIU) August 2021

- Design TxFuse based on the FDMI architecture from Seagate's storage system, a novel architecture that allows developing different storage features as plugins to improve CORTX distributed storage system. TxFuse enables the development and deployment of features independently of the storage system layer.
- Define a taxonomy of plugins that uses transactional coupling and a reliable notification mechanism. We identify three classes of plugins, class A, class B, and class C, which differ in how they interact with the storage system. Plugins execute code in response to client-initiated operations on the storage system.
- Simplify integration and deployment among distributed nodes using docker containers.
- Evaluate different plugin prototypes based on performance and development complexity. We count lines of code (LOC) and calculate cyclomatic complexity (CC) per feature in plugins and MinIO's native features.

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.**

Societies & Activities

- Planner, presenter, and team leader at FIU for 3 projects that include collaborations with globally distributed members from the technology industry and from universities in **2018, 2019, 2020, 2021, 2022,** and **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**.
- Participant in UPsilon Pi Epsilon, UPE activities on campus in **2019, 2021,** and **2022**.
- Volunteer for the Google CS First program as a Coding instructor for kids under the age of 8 teaching general computer science concepts, and critical thinking problems using Scratch in **2019**.