Research Statement

Janki Bhimani, Assistant Professor,
Knight Foundation School of Computing and Information Science, Florida International University
jbhimani@fiu.edu, bhimanijanki@gmail.com • Linkedin • Website • Google Scholar • Mobile: +1(857)991-9868

(*Please visit my* Website *for the most updated information.*)

[Overview.] The global datasphere is projected to reach 163 zettabytes by 2025, indicating a profound shift in our daily lives driven by data. As data centers adapt to meet escalating computational demands, the challenge of managing this abundance of information becomes critical. My decade-and-a-half-long research focuses on **pioneering disruptive solutions for memory management and storage systems to address this challenge**. The goal is to establish robust, parallel, and reliable computing systems capable of thriving in our data-intensive landscape. This work not only aligns with current federal funding priorities but is imperative for shaping the future, where innovation and efficiency converge to redefine the boundaries of large-scale data-intensive workloads.

In my research, I have achieved significant milestones by securing substantial funding of **$4 million**, with **$1.1 million** as Principal Investigator from major sources such as **NSF**, **Samsung Semiconductors**, and **Cyber Florida**. Notable, I am proud of my recent NSF CISE Core small award as sole PI focusing on memory management with machine learning (ML). Currently, I am awaiting decisions on **five proposals**, two of which I lead. Looking forward, my plans involve pursuing grants such as DOE CAREER, NSF Core Medium, and AFOSR to further advance my research. I have published **35 conference** and **10 journal articles**, with **8 patents** to my name and a notable **1000 citations**. I have been honored with prestigious awards, including **FIU Top Scholar**, **KFSCIS Excellence in Applied Research**, **Distinguished Reviewer**, and **Best Paper Awards**, acknowledging the impact of my work. My commitment to mentoring is reflected in the successful graduation of **one Ph.D. student** who now serves as an Assistant Professor. Over the past five years, I have mentored a diverse group, including eight Hispanics, one Asian American, and three women. Currently, I am supervising a dynamic team comprising **three Ph.D.**, **three M.S.**, and **three undergraduate** students. Establishing fruitful collaborations is integral to my research approach. I have forged partnerships with esteemed institutions such as the University of Maryland, University of Chicago, Argonne National Lab, and Syracuse University. Additionally, I have collaborated with industry leaders **Samsung Semiconductors** and **IBM Research**, contributing to the advancement of our collective research goals.

[Research Experience and Successes.] My research philosophy centers on incremental advancements towards "end-to-end system design" to tackle societal challenges. I thrive on projects involving emerging technologies, engaging in interdisciplinary research, and viewing theoretical investigations as tools for practical improvements. My research has focused on enhancing two important components of computer systems: Memory and Storage. My research contributions mainly in last five years as a faculty member can be classified mainly into the following three directions.

I/O Behavior Modeling, Performance Prediction and Optimization: In the era of "Big Data," where multiple data processing applications coexist in data centers, I/O activities exhibit significant variations. The configuration of existing storage systems, typically done during installation and then permanently maintained, is becoming insufficient. Modern data processing systems present three key challenges. First, simultaneous operations from multiple applications create interference, impacting the performance of solid-state drives (SSDs), unlike hard disk drives (HDDs) with limited bandwidth. Second, while HDD reliability relies on internal mechanical components, SSDs are sensitive to user and operating system I/O workloads, necessitating optimal configuration for persistent storage. Third, diverse SSD types, including multi-stream SSDs and Key-Value (KV) SSDs, come with distinct internal algorithms and parameters. Selecting and tuning these algorithms based on I/O activities is crucial for optimal performance and flash endurance. Therefore, the research objective is to **model the complex I/O activities of diverse applications, dynamically tuning the internal algorithm parameters of flash-based SSDs for optimal performance and reliability.**

Learning and Management in Tiered Memory Systems: Further, in the past decade, machine learning (ML) has undergone astounding growth, permeating various industries, including storage. Recognizing the increasing significance of ML in tandem with the proliferation of big data and the expanding storage market, there is a paramount need to optimize storage systems. To address the challenges posed by vast data and enhance memory accesses, tiered memory systems are gaining popularity. These systems employ high-speed memory like DRAM for frequently accessed upper-tier data and slower but larger memories like NVMe NAND flash, 3D-Xpoint, and CXL memories for lower tiers. These systems prove crucial for efficient data management, contributing to improved performance and efficiency, reducing data access times, and lowering overall computing costs. Thus, we **design novel tiered memory management techniques, leveraging ML's power and addressing its limits and overheads as a versatile solution to enhance various aspects, including scanning, migration, allocation, parameter tuning, and task scheduling, ultimately optimizing performance, Quality of Service (QoS), and resource utilization in in-memory databases and analytic frameworks.**

Towards Efficient In-storage Indexing and Device Endurance: In the domain of in-storage indexing, a concept originating in the 1990s, the practical implementation has encountered obstacles, as exemplified by Seagate's 2014 endeavor with Kinetic HDDs, constrained by HDD I/O limitations. However, the prospect of efficient in-storage indexing has gained momentum with flash-based SSDs. Early efforts, including our preliminary research, demonstrate that the combination of a fast and lightweight KV database or POSIX-compliant file system with a key-value SSD (KVSSD) performing in-storage indexing outperforms traditional block SSDs. The development of Efficient In-Storage Indexing Devices (ISIDs) poses challenges that necessitate attention to ensure optimal performance and functionality.

Firstly, the role of storage device models is pivotal in computer systems research, addressing research gaps related to performance analysis, algorithm development, system evaluation, resource management, and realistic simulations. The lack of a low-cost open-source research platform hampers rapid adoption. Secondly, the design of ISIDs for diverse workloads demands meticulous consideration of indexing techniques, query optimization, data access patterns, and data distribution within the constraints of limited device resources. Therefore, we **develop ISID models that capture internal feature dependencies and support dynamic model calibration. This aims to develop new index management techniques efficiently utilizing limited on-device resources while considering flash-specific constraints to optimize endurance and latency for a multi-tenant environment.**

In conjunction with above my team is currently pursuing various other cutting-edge projects. We challenge the assumption that Solid-State Drives (SSDs) are less susceptible to vibration than Hard Disk Drives (HDDs). Extensive testing reveals that, even within specified limits, vibration significantly impacts SSD I/O performance, with short-term vibrations exhibiting lingering effects and long-term exposure leading to over **30%** performance degradation. This has crucial implications for autonomous vehicles' safety and raises concerns in data center performance, affecting tail latency and Service Level Agreements (SLAs). Our findings and anonymized experimental data are publicly available. Building on these insights, our research extends to the automotive industry's shift to autonomous driving, emphasizing temperature's role in the reliability of electronic control systems (ECS). Operating under harsh conditions and increasing power density in ECS integrated circuits present challenges. Our study maps automotive applications onto multiple-core processing architectures, considering temperature and system-level reliability. Employing a mathematical programming model and a genetic algorithm, we optimize system-wide reliability, ensuring peak ECS operating temperature. Computationally efficient techniques for system-wide mean-time-to-failure (MTTF) computation exhibit significant speed-ups, enhancing reliability analysis for automotive applications.

Another initiative involves critically examining wear leveling in SSDs, addressing challenges, assessing effectiveness, and advocating for capacity variance. Additionally, our team is dedicated to enhancing data storage and ingestion pipelines for ML training. We investigate novel approaches to Deep Neural Network (DNN) checkpointing beyond traditional file I/O methods, leveraging data science to revolutionize decision-making in storage I/O. Lastly, our commitment extends to advocating for positive change in educational systems through the project "Voices for Organizing Change in Educational Systems." These diverse undertakings showcase our comprehensive engagement in advancing knowledge and addressing pivotal challenges across multiple facets of data storage and computing.

[Publication Decisions.] Dedicated to advancing data storage and management, my 45 publications, 20 from FIU and 25 predating my affiliation, span journal articles and conference papers. Focused on Tier 1 journals like IEEE Transactions on Cloud Computing and ACM Transactions on Storage, my research covers diverse storage challenges, from dynamic memory allocation for Apache Spark to flash endurance in data centers. For rapid dissemination, I publish in conferences like IEEE CLOUD and ACM HotStorage, with Best Paper Awards highlighting impact. Google Scholar shows 1000 citations, an h-index of 16, and an i10-index of 20, reflecting broad influence. Beyond publications, talks at conferences and institutions like IBM Research and Samsung demonstrate my commitment to disseminating insights and bridging academia with industry.

[Research Funding.] In the last five years at FIU, my research endeavors have been supported by multiple NSF grants and other funding opportunities, including the Global Research Outreach (GRO) Program by Samsung and Cyber Florida. I have been fortunate to secure a total of eleven grants valued over $4 million, with eight of them as the lead/sole Principal Investigator. Notably, I am proud of my recent highly competitive NSF CISE Core small award, where I am the sole PI. This award focuses on designing new memory management techniques for in-memory databases and analytic frameworks, leveraging the power of machine learning (ML). This funding underscores the significance of my work in the national discourse and allows me to explore the design of sustainable ML-based tiered memory systems, considering trade-offs between performance, Quality of Service (QoS), and resource utilization using different ML models for various design decisions. Currently, I have five proposals under review, with two of them as the Principal Investigator. Looking ahead, I plan to apply for prestigious grants, including the DOE CAREER award, NSF Core Medium, and AFOSR grant in the upcoming year.

[Future Plan.] In my envisioned trajectory, I am committed to further solidifying my standing as an expert in memory management and storage systems, with a dedicated focus on emerging technologies. The essence of my work will persist in actively contributing to and influencing the national discourse surrounding efficient, reliable, and enduring data management techniques. Each successive phase of my research is strategically crafted to make practical and measurable strides toward diversifying the landscape of optimal storage techniques. My future plans encompass an extension of recruitment strategies to forge partnerships with institutions, aiming to intentionally diversify my doctoral student cohort to mirror the rich diversity present in the FIU undergraduate student population. Simultaneously, my dedication to securing funding remains unwavering, spanning sources such as the FIU Foundation (industry), government agencies, and the National Science Foundation. Looking forward, I aspire to leverage the national network I have cultivated to foster collaborations on a broader scale. Initiating endeavors for large-scale, multi-institutional funding is a pivotal goal, intended to amplify the impact of my research endeavors and revolutionize storage practices.