



Edward E. Whitacre Jr.
College of Engineering™

Computer Science

New Storage Technologies for Big Data Processing on Cloud and Datacenter Infrastructures

Janki Bhimani, Ph.D.
Prospective Faculty Candidate
Northeastern University
Thursday, February 7, 2019
2:30 p.m.
ECE, Room 217

Abstract

A new era of “Data Age” is approaching today. Data is the fuel for analytics of all the emerging technologies of Internet-of-Things (IoT) and cloud computing. Data management plays a critical role in delivering real-world impacts. However, it is challenging for any systems to efficiently manage data to achieve low latency, high throughput, and good endurance. A good ecosystem is highly demanded to provide fine coordination among multiple facets of the system such as parallel computing, hierarchical memory caching, and low I/O latency storage. Mitigating bottlenecks at each facet/layer is essential for accelerating overall production-scale deployments. I will present my research that aims to develop new data management techniques and schemes for building an efficient and reliable computing system for data-intensive applications. I will introduce my major research contributions related to three data management components and overview cutting-edge technologies such as containerized virtualization, scalable big-data infrastructures such as Spark, and new flash-based technologies such as multi-stream and key-value Solid State Drives (SSDs). I focus on studying resource management for cloud computing and datacenters and investigating big-data and high-performance computing platforms. I will present new insights and challenges that are posed by modern evolving data processing applications and new storage technologies, as well as, our proposed approaches for data management to improve overall endurance of emerging large-scale storage systems. Finally, I will discuss the challenges in efficient and reliable management of upcoming heterogeneous data-centric computing requirements, and possible approaches toward developing infrastructures to accommodate these future requirements.

Biography

Janki Bhimani is a fifth-year Ph.D. student from Department of Electrical and Computer Engineering at Northeastern University. She received her B.S. in Electrical and Electronics Engineering from GITAM University in 2013 and her M.S. in Electrical and Computer Engineering from Northeastern University in 2014. Her primary research interests are Flash-Based Storage Enhancement, Big Data Processing, Cloud Computing, and Parallel and Distributed Computing. She received two Best Paper Awards from IEEE CLOUD 2018 and IPCCC 2017. She published 19 peer-reviewed conference papers since 2013, and among them, 10 papers are in highly selective conferences such as IEEE CLOUD, MSST, and IPCCC. She also published 5 journal papers, two of which are in IEEE Transactions. She has worked at Samsung Research Labs (MSL and MPL) as summer interns for three times and filed six top graded patents during her internships. In Fall 2017, she also worked as an instructor at Northeastern University, teaching an undergraduate core course “Fundamentals of Engineering Algorithms.” She received excellent feedback from her students with a high rate of 4.4/5 for her overall instructor effectiveness. In her free time, Janki is a creative visual artist. Far from home, amidst nature she finds her inspiration to paint. She enjoys understanding the impact of art on human psychology, and she painterly bring motivation, healing, and encouragement through the canvas.