Department of Advanced Design
and Systems Engineering
香港城市大學
City University of Hong Kong

Reconfigurable Compute-In-Memory on Ferroelectric Memory Technology: A Cross-Layer Research

27 Feb 2023 (Mon) | 10:30 am

Seminar Link: https://cityu.zoom.us/j/99056682896

Mr. LIU Xiwen
PhD candidate
Department of Electrical
and System Engineering
University of Pennsylvania, USA

Abstract

Current computing systems are mainly constructed on the von Neumann architecture, where data needs to be transferred to a processing unit from memory components. The latency associated with accessing data from the memory units is a key performance bottleneck for a range of data-intensive applications in the convergence of big data and AI. Several solutions have been proposed to mitigate and overcome this bottleneck, with a prominent one being placing memory and logic units in close physical proximity. While significant progress has been made along those lines at both technology and architecture levels, a transformative approach would be to perform computing functions precisely where the data are stored using memory devices. This is known as compute-in-memory.

In this talk, I will begin by discussing the most recent advancements in the CMOS-compatible ferroelectric memory technologies on aluminum nitride platform. Second, I will present a reconfigurable compute-In-memory system on field-programmable ferroelectric diodes, allowing for on-chip memory, parallel search, and neural network operation. Last, I will discuss the conceptualization and demonstration of a programmable parallel search architecture - analog content-addressable memory (ACAM) on complementary Si-CMOS ferroelectric field-effect-transistor memory. The deployment and acceleration of attention-based deep neural network on ACAM will also be presented.

About the Speaker

Mr. Xiwen Liu is currently a PhD candidate in Electrical and System Engineering at University of Pennsylvania, along with pursuing a research master's degree in Statistics at the Wharton School. His research interest includes non-volatile memory technology, compute-in-memory system, and technology-circuit-algorithm co-design for AI accelerator. His work on memory-enhanced computation have been highlighted by many media sites and recognized by the Bell Lab Prize 2022.

Enquiry: 3442 8422 | All are welcome