IDEC Chip Design Contest



An eFlash-Based Computation-in-Memory for Al Edge Computing

Injun Choi, Jongyoon Choi, and Minkyu Je

Laboratory of Integrated Microsystems Powered with Advanced Circuit Technology,
School of Electrical Engineering, KAIST

SkB 32kB 1MB DRAM Memory Size (D) Memory access energy

CPU
Control Unit

Many Data
Movement
Storage
Cells

CPU
Control Unit
Logic Unit

CIM Architecture

Memory
Storage
Cells

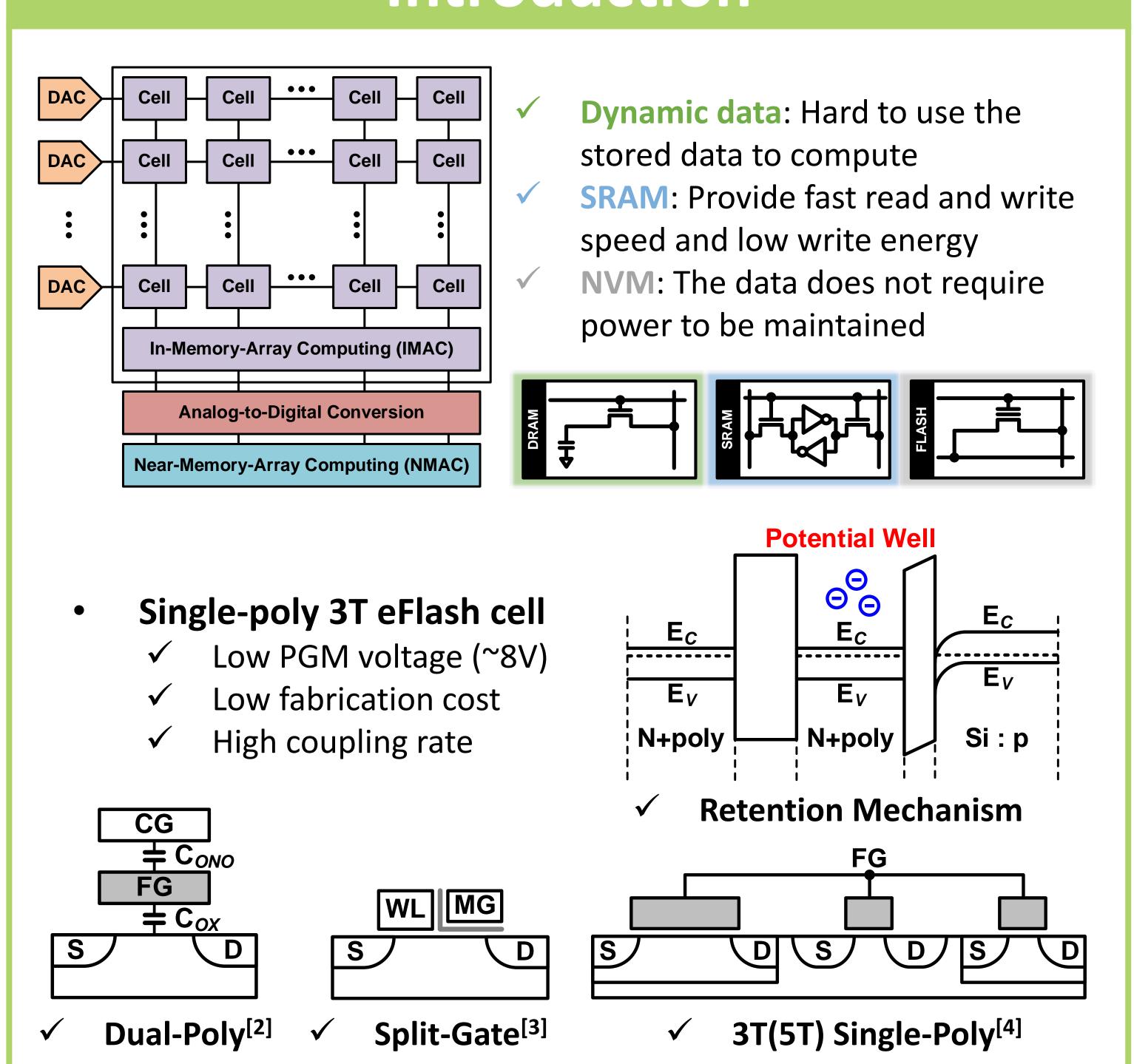
Logic Unit

→ More energy-efficient

- To achieve energy-efficient AI computing
 - ✓ It is crucial to minimize the amount of data movement
- Memory access energy is greater than computing energy
 - ✓ Computation-in-Memory(CIM) can drastically reduce data movement and implement a more energy-efficient system
 - ✓ CIM is a memory-centered computing architecture that integrates computing units in memory
- CIM enables the implementation of energy-efficient systems
 - ✓ Reduce the movement of data
 - Effective for parallel processing

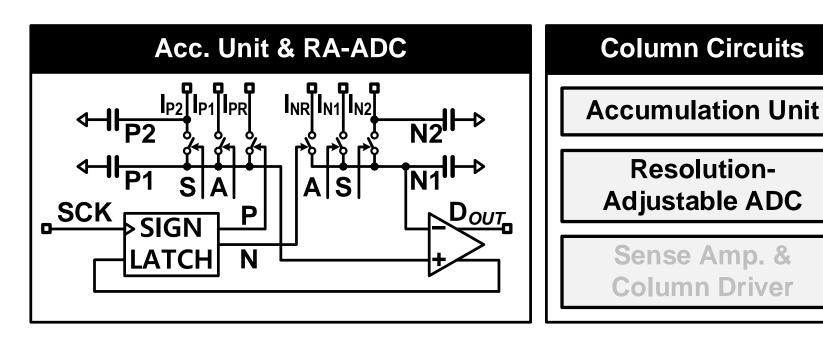
Introduction

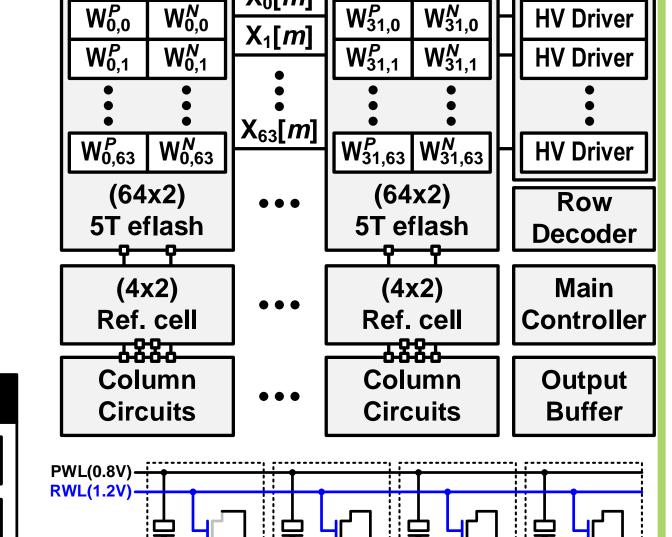
>> Computation energy^[1]



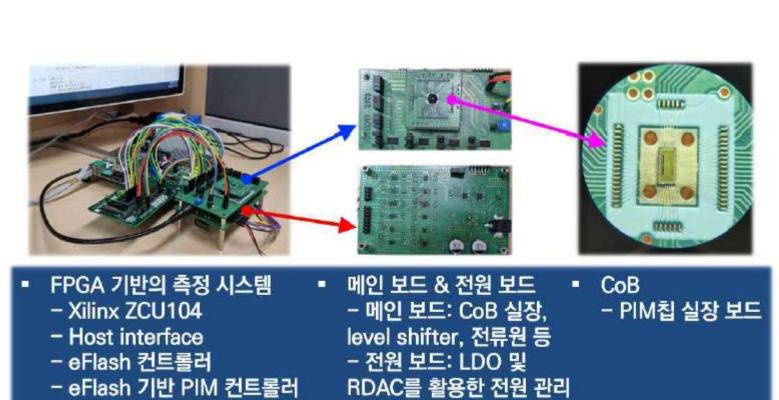
Implementation & Measurement

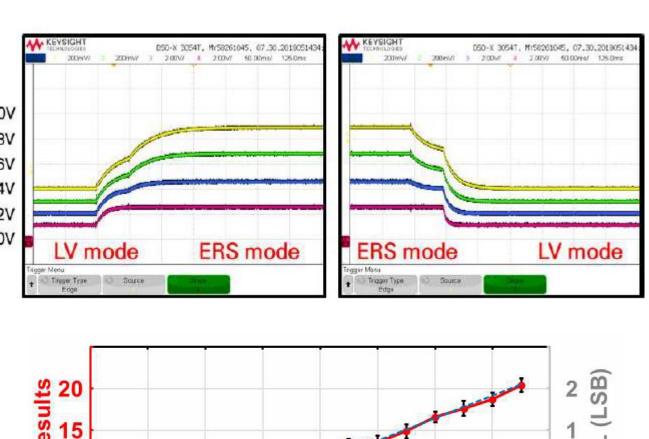
- High-level block diagram
 - ✓ Multi-level eflash pair
 - ✓ Biasing for wide-swing BL
 - ✓ Charge domain computing
 - ✓ Resolution-adjustable ADC





Measurement results





Conclusion & Acknowledgement

- Challenges
 - ✓ BL Limited dynamic range and nonlinearity issues
 - ✓ ADC and DAC overhead
- Achievements
 - ✓ Low-cost charge domain computing circuit and ADC
 - ✓ Analyze MAC values for a sufficient margin

References

[1] M. Horowitz, ISSCC 2014

[2] H. Kojima, IEDM 2007

[3] J. Yater, IMW 2009

[4] S.-H. Song, JSSC 2013

- Acknowledgement
 - The chip fabrication and EDA tool were supported by the IC Design Education Center(IDEC), Korea.

