IDEC Chip Design Contest

eDRAM based digital computing in memory neural network accelerator with enhanced retention time

[1]Inhwan Lee, [1]Hyunmyung Oh , [2]Jehun Lee, [2]Jaeyong Jang and [2]Jae-Joon Kim

[1] Pohang university of science and technology (POSTECH)

[2] Seoul National University

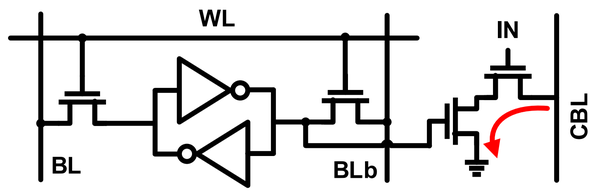
Background

Motivations of eDRAM based NN accelerator

*CIM: Computing-in-Memory, PVT: Process, Voltage, Temperature

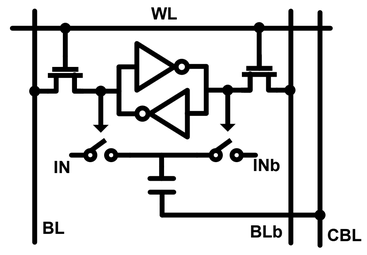
- SRAM cell based CIM design

Resistive Cell



Non-linear operation, instability due to PVT variation

Capacitive Cell



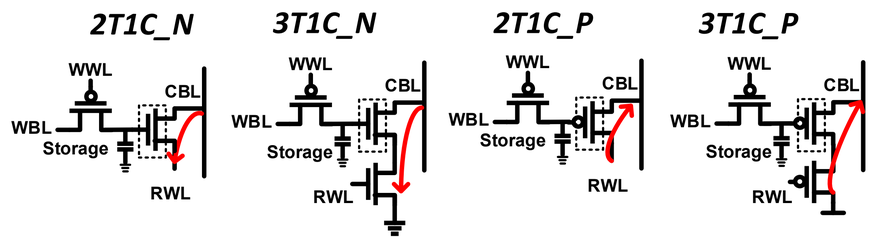
Linear operation, stability from PVT variation

SRAM cells have large cell sizes that limit their memory capacity

There is a need for smaller cell sizes to overcome the issue of limited memory capacity.

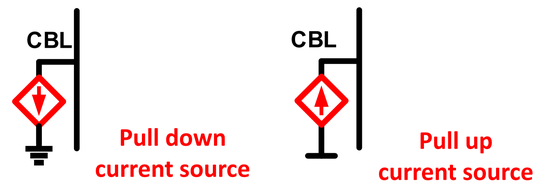
Previous research

- eDRAM cell based CIM design



- Modeling of eDRAM cell designs

Current source ∝ V(storage_node)



Pull down current source

Pull up current source

Advantage

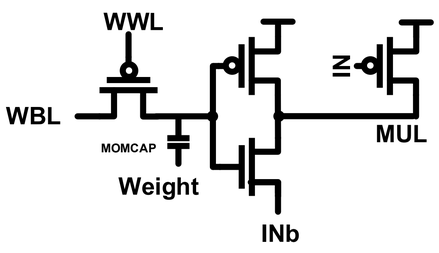
- ✓ Utilize 2 to 4 transistors to implement CIM operation with a smaller cell size.
- ✓ Logic compatible with CMOS process

Disadvantage

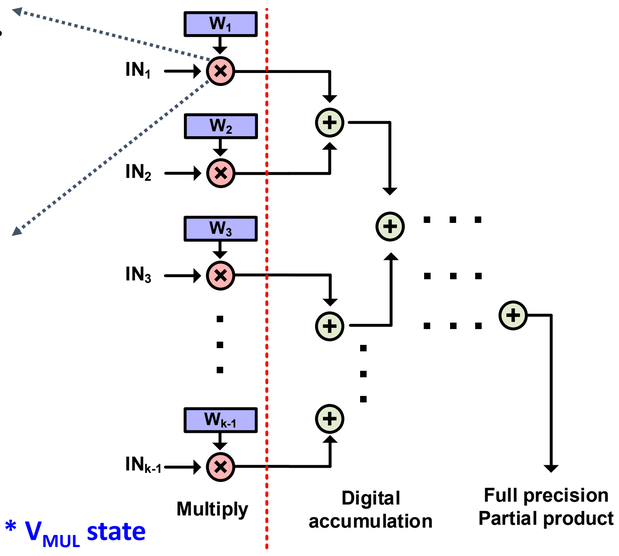
- ✓ Currents are vulnerable to PVT variation
- ✓ Unstable output value due to leakage current
- ✓ Short retention time

Proposed eDRAM array

Proposed cell and architecture



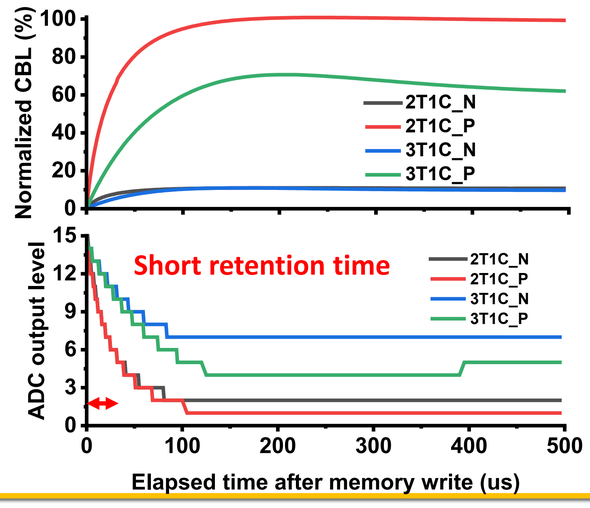
IN \ W	0	1
0	1	1
1	1	0



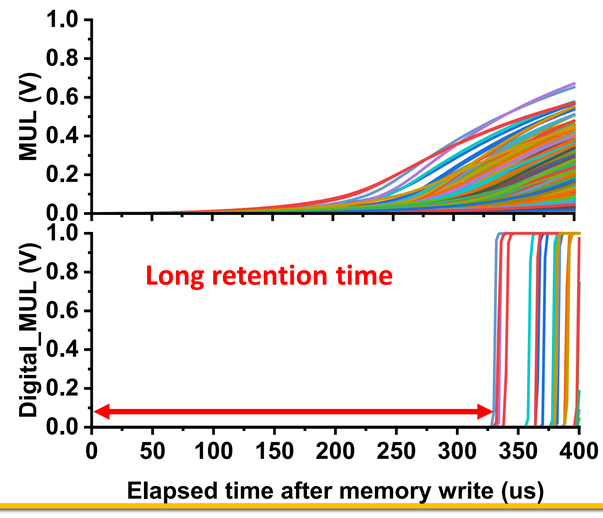
* V_{MUL} state

Increased retention time

- eDRAM Analog CIM

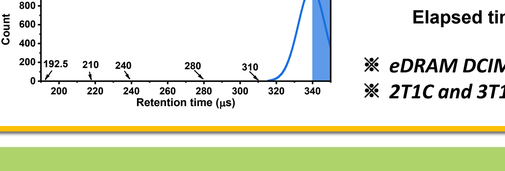


- eDRAM Digital CIM

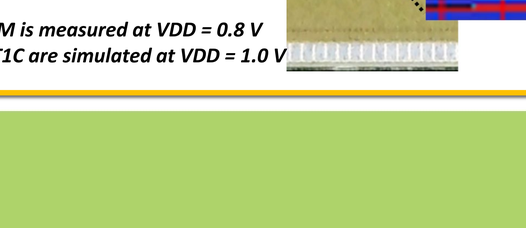


Evaluation

CIM operation



Throughput and energy consumption



※ *We assume array-level (array size = 256x64) operation for 2nd CNN layer*

※ *ACIM supply voltage = 1.0 V , DCIM supply voltage = 1.0 V for Core cell and 0.7 V for Peripheral circuit.*

Retention time

- Measured heatmap

- Neural Network accuracy with time

※ *eDRAM DCIM is measured at VDD = 0.8 V*

※ *2T1C and 3T1C are simulated at VDD = 1.0 V*

Die photo and layout of eDRAM array

Technology	28 nm CMOS
Supply voltage (V)	1.0
Frequency (MHz)	70
Input / Weight / Output	4/4/ 12
Array size	256x4 (1kb)
Total Area	190.7 um x 227.25 um

Conclusion

- A novel eDRAM cell mimicking a digital logic gate for digital CIM operation is proposed
- Long retention time and accurate MAC values are accomplished due to DCIM operation
- x 15 increased throughput and x 1.7 reduced energy consumption is obtained due to enhanced retention time.

Acknowledgement

The EDA tool was supported by the IC Design Education Center (IDEC), South Korea.

반도체설계교육센터 IC DESIGN EDUCATION CENTER