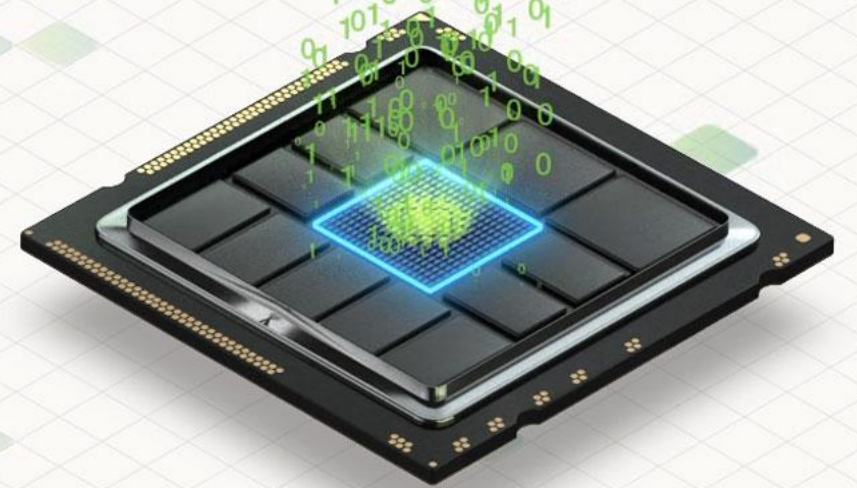


ReRAM Gets Real: From Concept to Market

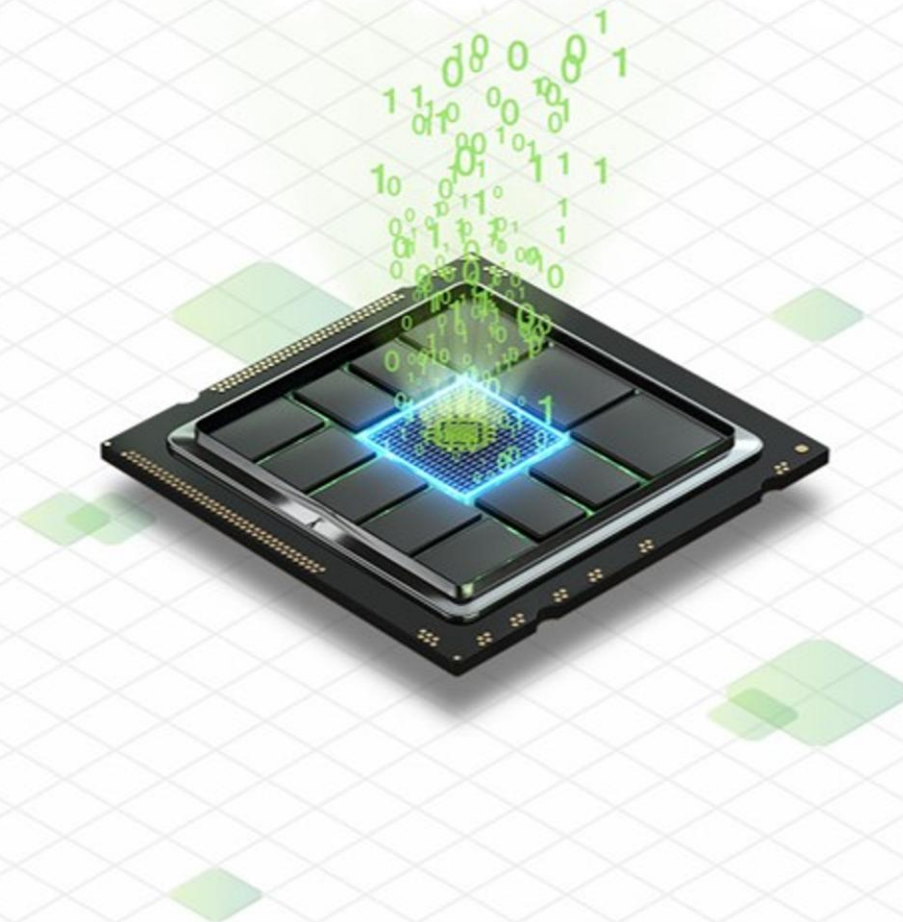

MPSoC'25

Gabriel Molas,
Chief Scientific Officer, Weebit Nano

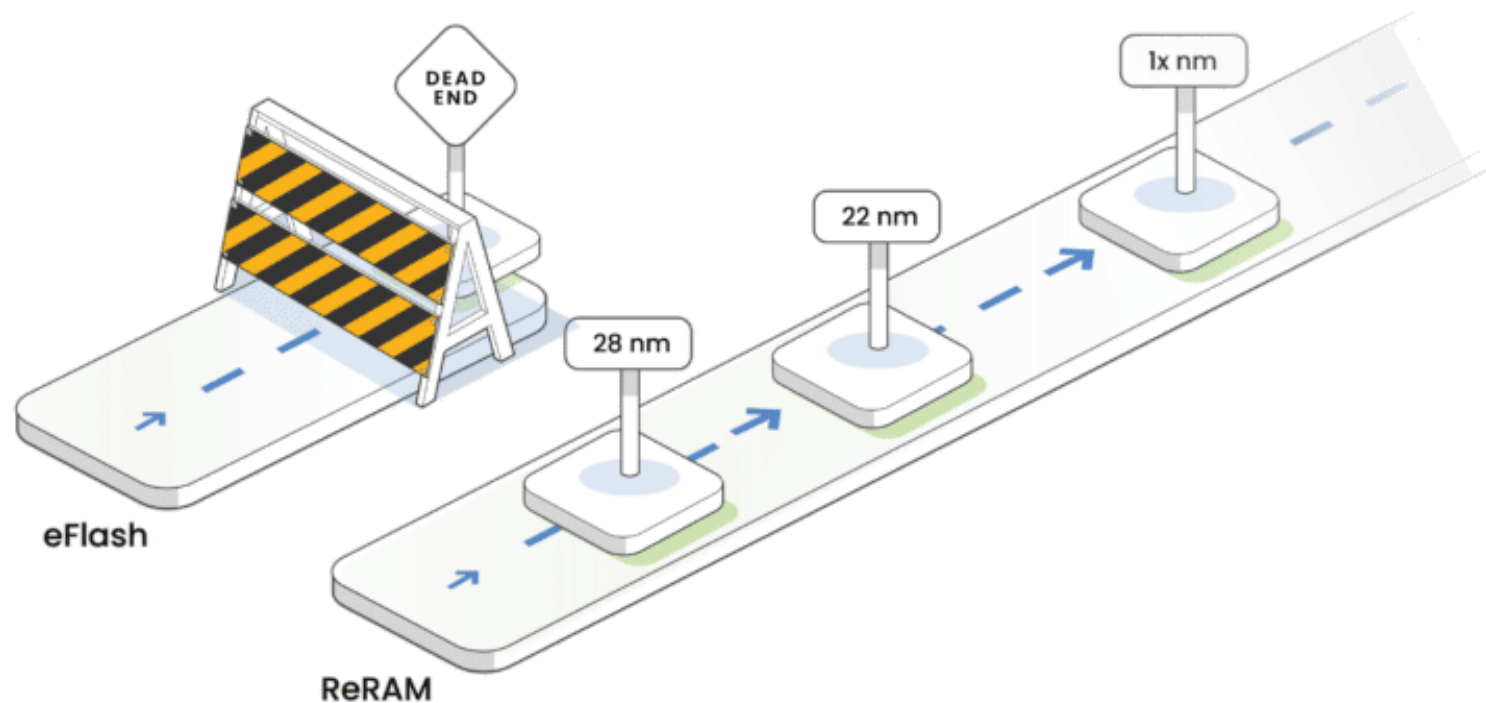
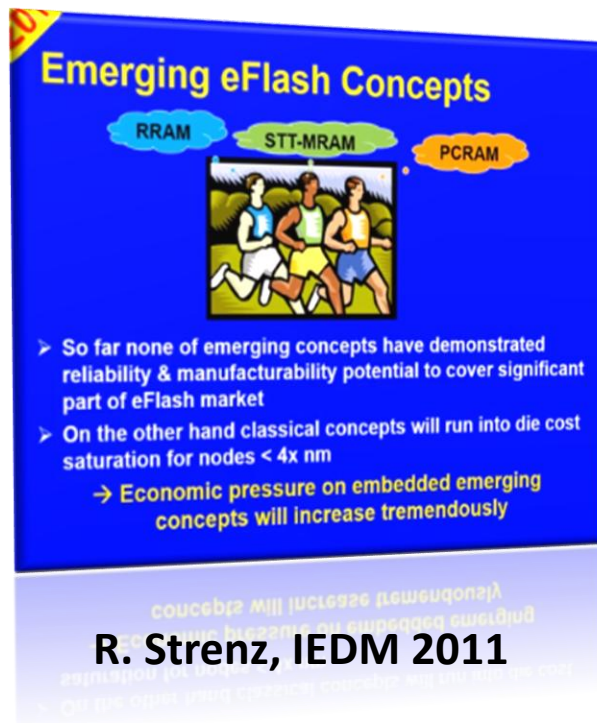


Outline

- ❖ **Context and introduction**
- ❖ Technology challenges and solutions
 - ◆ Bit error rate
 - ◆ Stability
- ❖ AI opportunities
- ❖ Conclusions

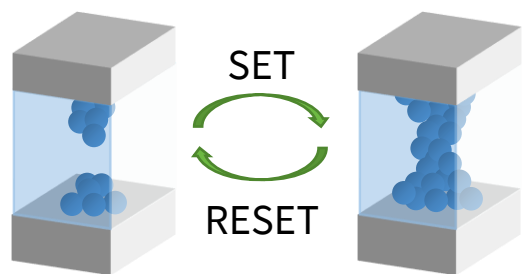


Emerging NVM for MCU



- ❖ Flash is stuck at 28-22nm because of the integration cost and complexity increase
- ❖ **Emerging NVM** (ReRAM, PCM, MRAM) went through long R&D and are now replacing Flash for embedded applications
 - ◆ The market is expected to grow significantly over the next years

ReRAM History



1960

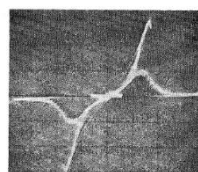
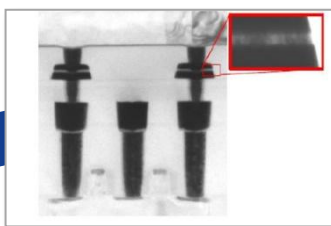


Fig. 1—Current-voltage locus, 60 cps. Ordinate scale is 40 mA/div, abscissa 2V/div. Superimposed is the high initial resistance locus resulting from random voltage removal at about 7-volts peak. The upper negative resistance curve is for voltage increasing, the lower for voltage decreasing.

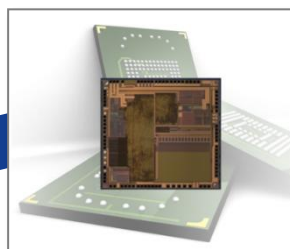
Resistive switching
phenomena in
oxides

2000s



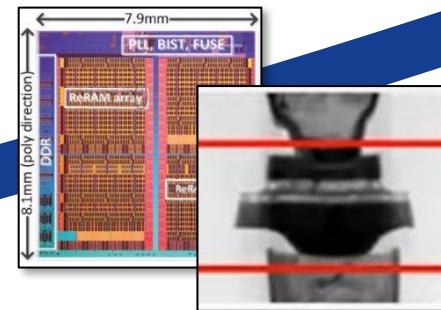
Samsung
RRAM integrated in
0,18μm...
40 years later

2010s



Panasonic
1st commercially available
implementation of RRAM
Healthcare, security equipment
or sensor processing applications

2019



22nm FinFET
Intel, TSMC
10⁴ cycles, 85°C 10ys
retention

2022-2024

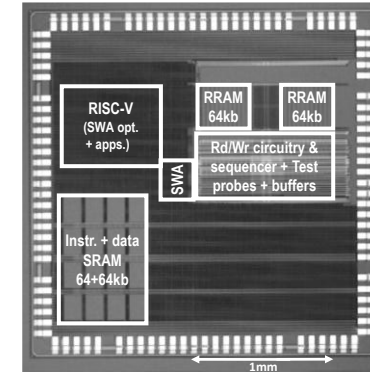


Infineon & TSMC
(Automotive μC
28nm)

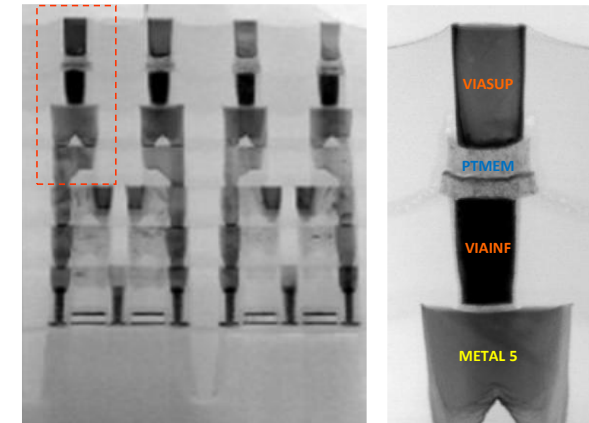
ReRAM products are now available for embedded applications,
50 years after the switching effect was discovered...

Weebit ReRAM

- ❖ Weebit has collaborated with CEA-Leti since 2016 to develop its ReRAM
 - ◆ Memory module fully functional and **qualified** for 150°C, 100K cycles
 - ◆ Mbit arrays demonstrated at **28nm**
 - ◆ Taped-out 8Mb prototype in **22nm FD-SOI**
- ❖ On the path towards mass production
 - ◆ Weebit ReRAM IP available for production in SkyWater PDK for customers
 - ◆ Testing 1Mb ReRAM module in DB HiTek's 130nm BCD process
 - ◆ Licensed ReRAM to onsemi, 65nm BCD process for automotive
- ❖ **R&D** continuously ongoing to improve technology and scale ReRAM technology to smaller nodes



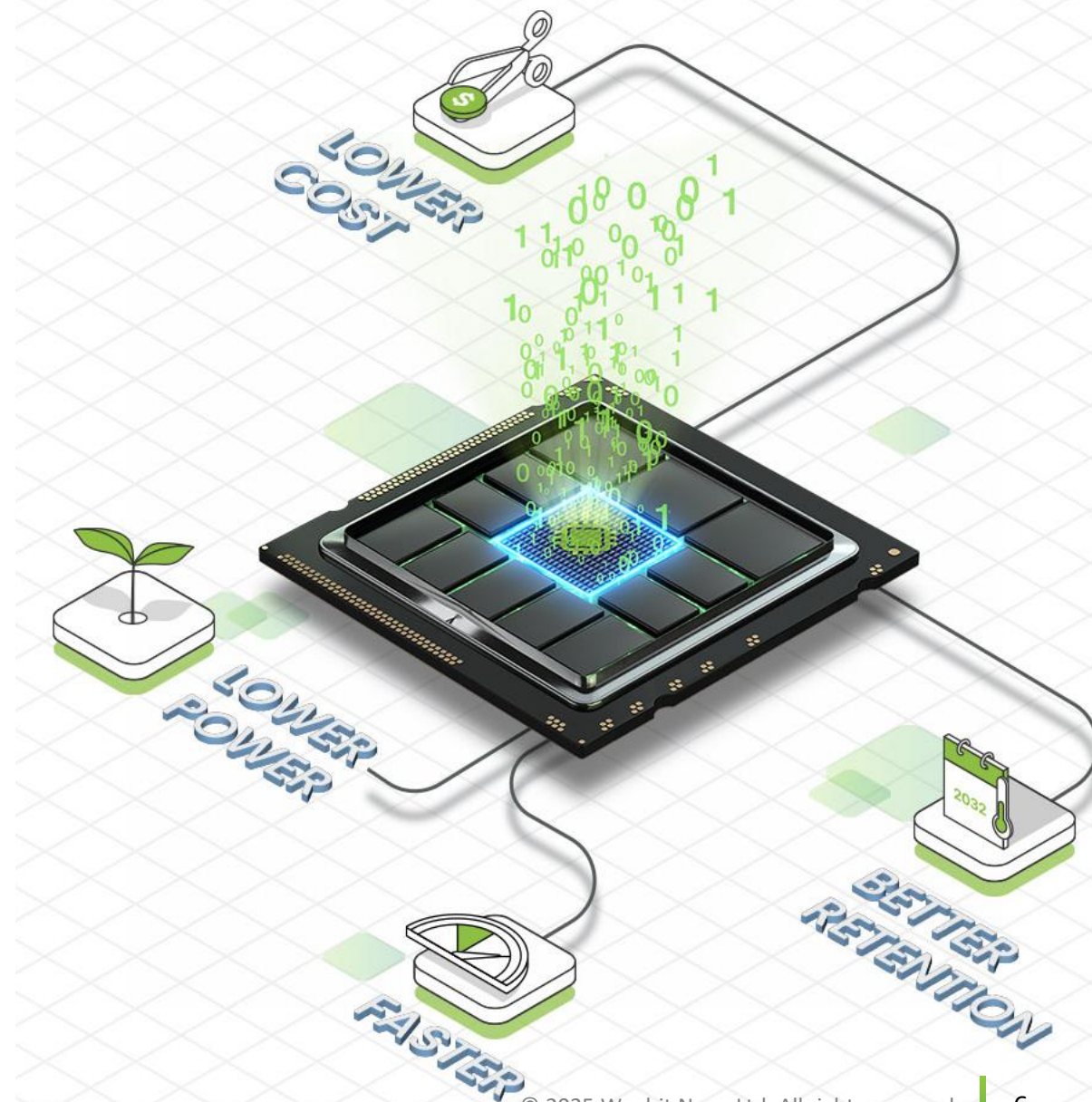
130nm ST/Leti MODULE



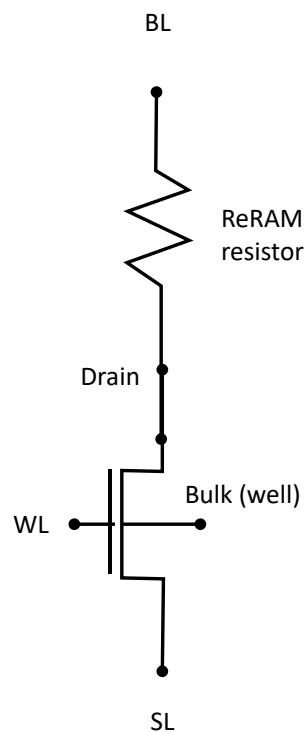
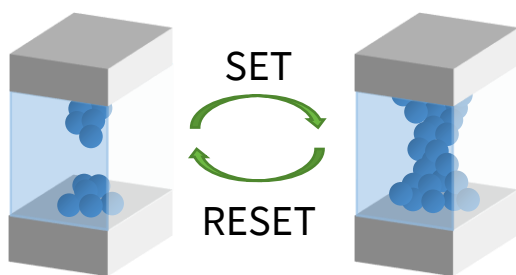
ReRAM in 28nm

Outline

- ❖ Context and introduction
- ❖ **Technology challenges and solutions**
 - ◆ Bit error rate
 - ◆ Stability
- ❖ AI opportunities
- ❖ Conclusions



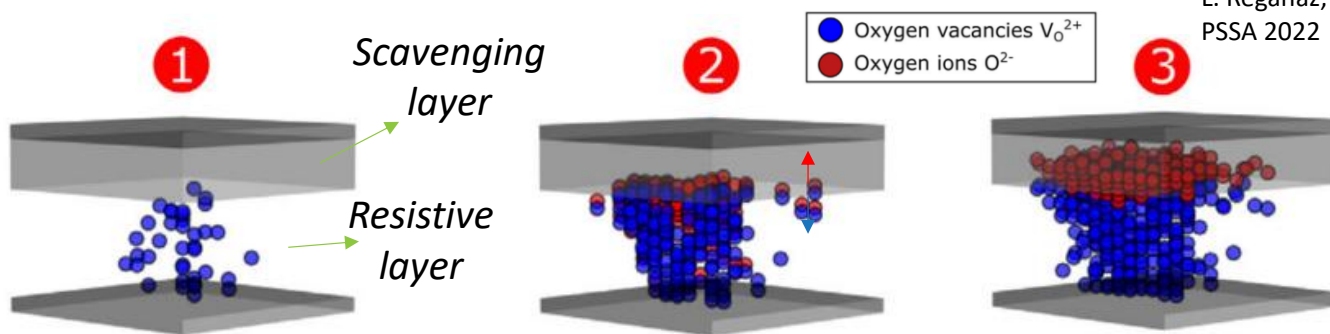
ReRAM Basics



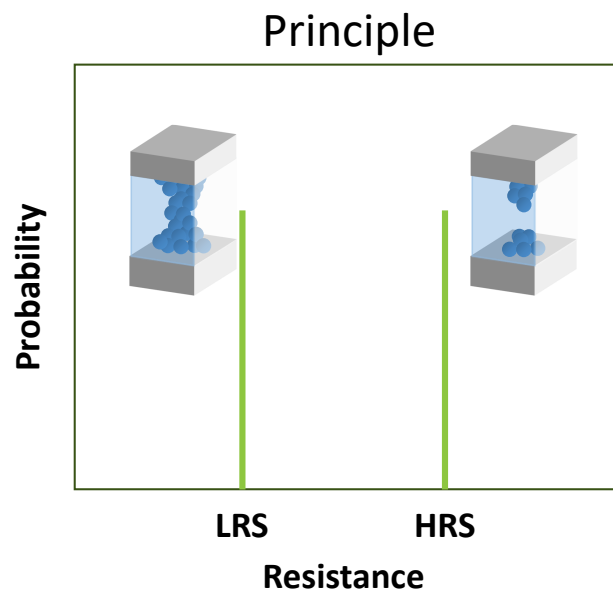
- ❖ **Operating principle**
 - ◆ Formation/disruption of a conductive filament in a resistive layer (metal oxide) – 10nm scalability demonstrated today (1D filament)
- ❖ **2-mask adder**
 - ◆ Very few added steps compared to other NVM technologies
 - ◆ Lower wafer cost than competing NVM technologies
- ❖ **Fab-friendly materials**
 - ◆ No contamination risk, No special handling, etc.
- ❖ **Using existing deposition techniques and tools**
 - ◆ Easy to integrate into any CMOS fab
- ❖ **BEOL technology**
 - ◆ Stack between any 2 metal layers
 - ◆ No interference with FEOL – Easier to embed with existing Analog / RF circuits
 - ◆ Easy to scale from one process variation to another
- ❖ **MCU/IoT: a natural fit for ReRAM**
 - ◆ Lower Power Consumption, Minimized Cost, High-level of Integration, good retention and endurance

Basics of ReRAM Switching Physics

L. Reganaz,
PSSA 2022



- ❖ Stochastic character of filament formation and dissolution
- ❖ ReRAM resistance is spread for HRS and LRS



Diffusion

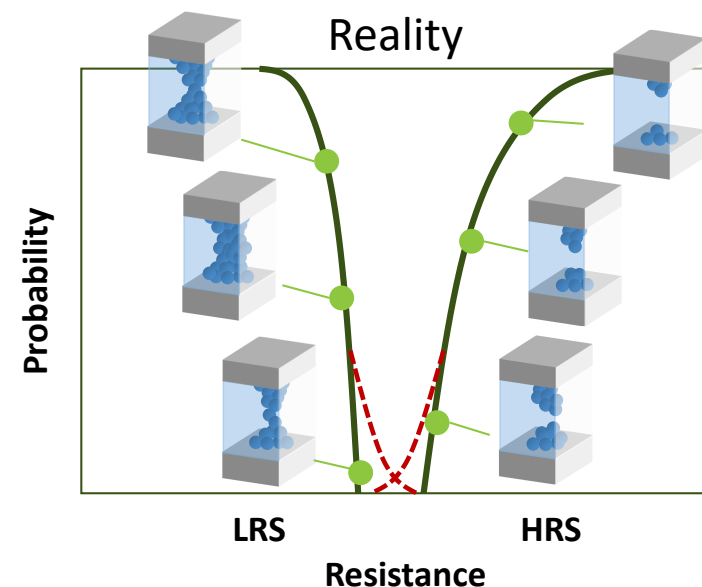
$$\Gamma_D(x, y, z) = \nu_0 \exp\left(-\frac{E_{A,D} - \frac{1}{2} Qd(1 + L_X)E}{k_B T(x, y, z)}\right)$$

Recombination

$$\Gamma_R(x, y, z) = \nu_0 \exp\left(-\frac{E_{A,R}}{k_B T(x, y, z)}\right)$$

Generation

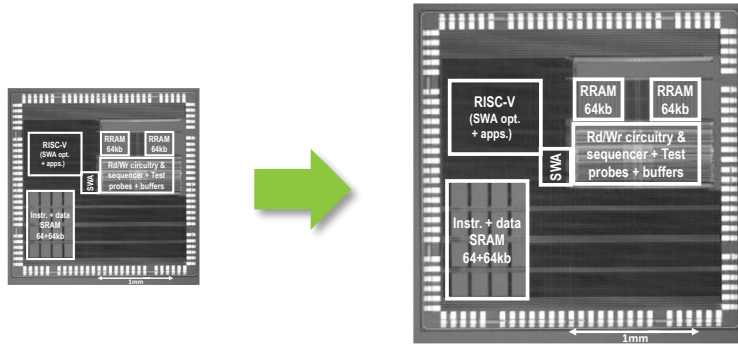
$$\Gamma_G(x, y, z) = \nu_0 \exp\left(-\frac{E_{A,G} - bE(x, y, z)}{k_B T(x, y, z)}\right)$$



- ❖ Generation and diffusion of VO_2^+/O_2^- pairs in resistive layer
- ❖ Oxygen ions absorbed in the scavenging layer (Top electrode acts as an oxygen reservoir)
- ❖ Conductive filament formed in the resistive layer when sufficient voltage is applied

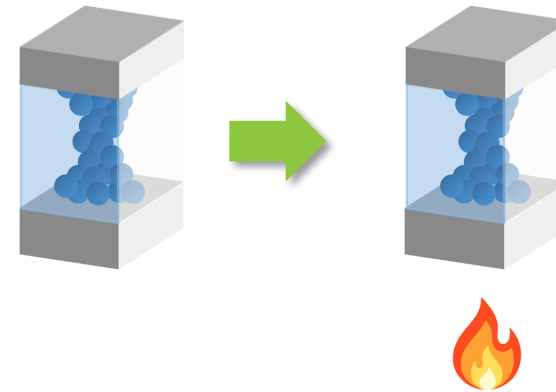
Technology Enhancements

- ❖ 2 important examples (among others...)



Margin / Error rate

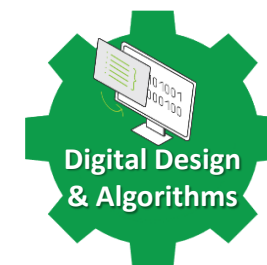
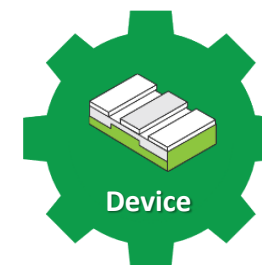
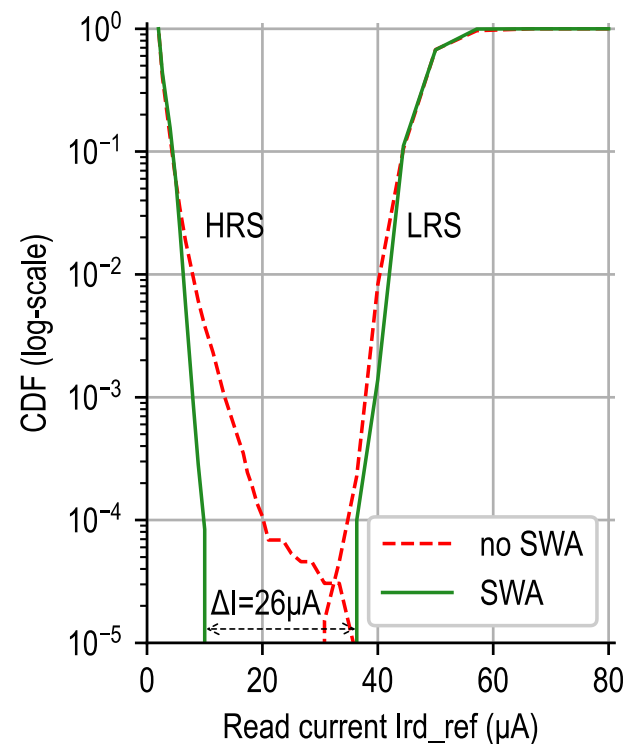
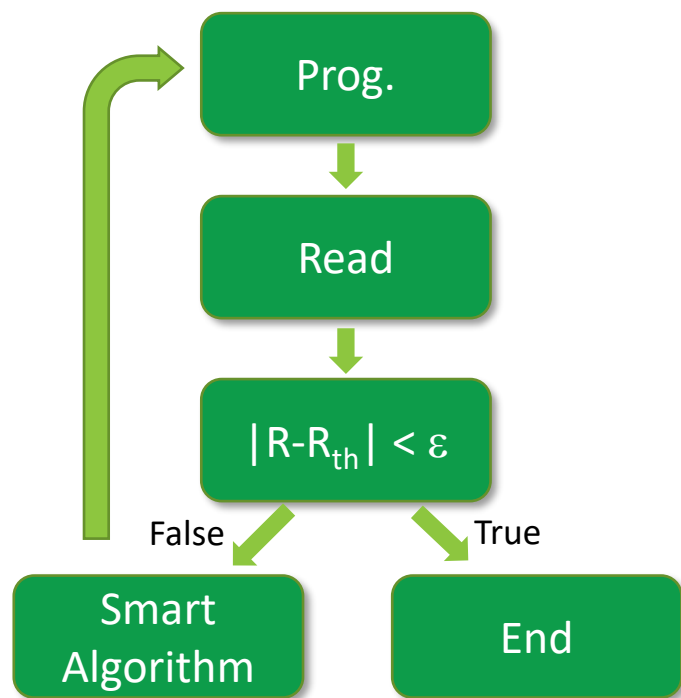
Lower error rate required as memory
capacity increases



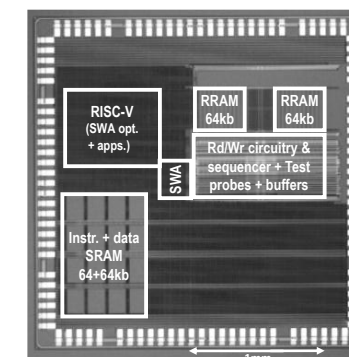
Stability

High ReRAM stability required for high
temp demanding applications
(automotive...)

Algorithms for Margin Increase



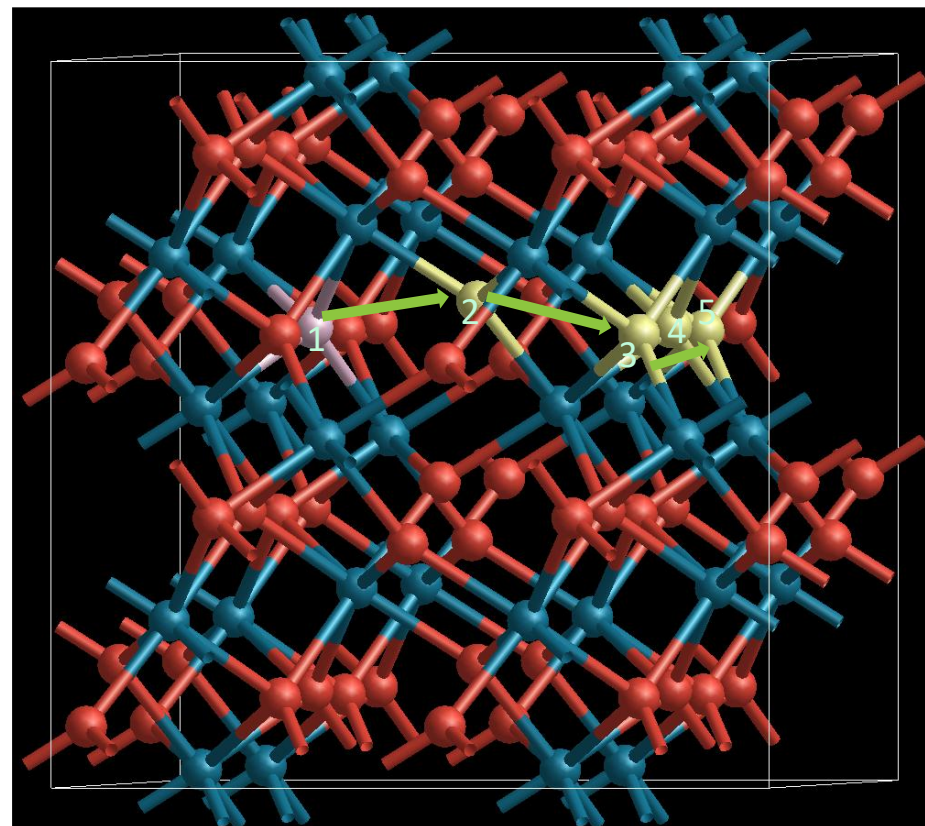
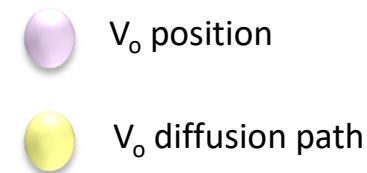
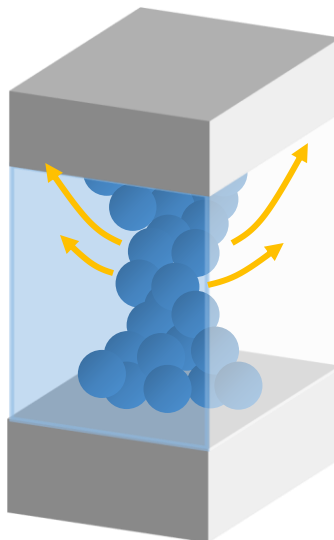
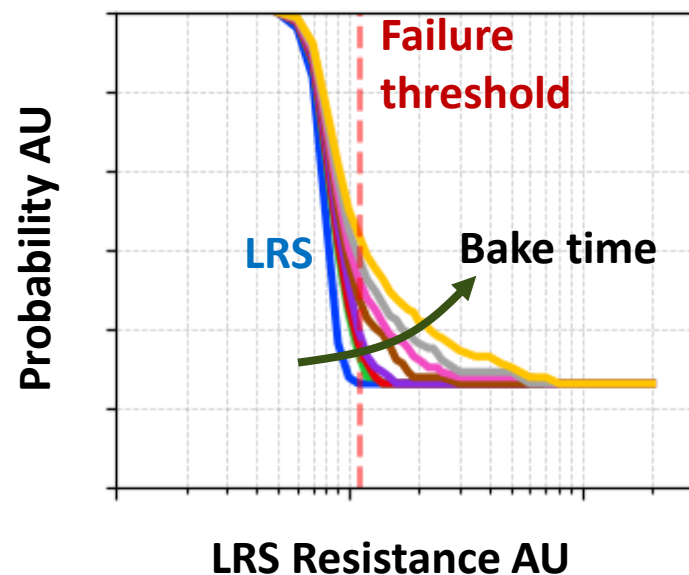
B. Giraud, IMW 2023
List, Leti – Weebit



- ❖ **Program algorithms:** resistance is read after programming, and ReRAM is programmed again if needed
- ❖ → Larger margin is achieved

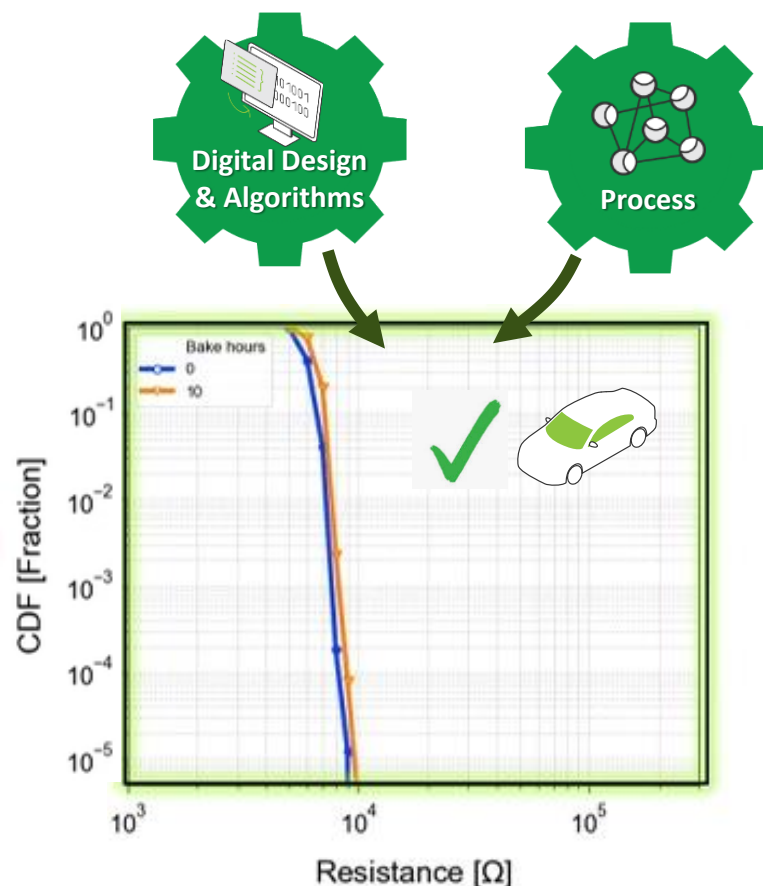
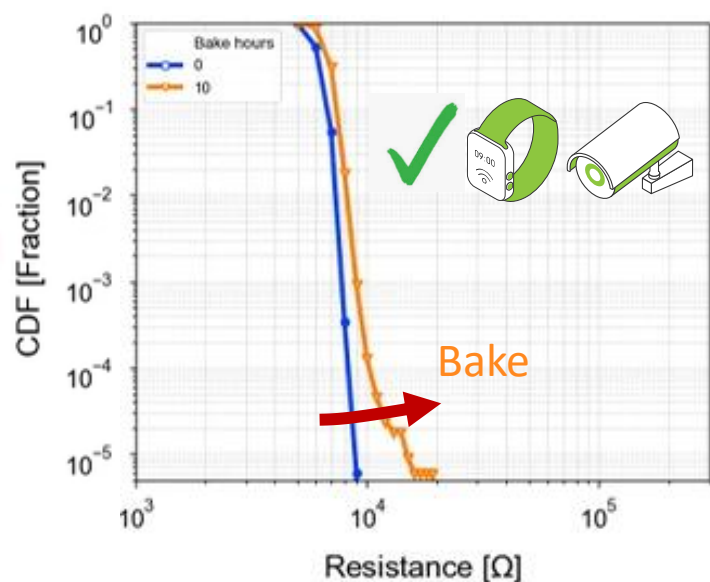
Basics of ReRAM Retention Degradation

Low Resistance State Retention responsible for data loss



- Some cells undersee V_O migration during high Temp. bake
- Progressive resistance distribution spread, leading to information data loss

Retention Improvement



- ❖ Improved retention
 - ◆ Part can be due to **more stable material** (lower oxygen migration energy)
 - ◆ Part can be due to **improved programming algorithm**, leading to more stable filament

- ❖ Stack optimization and new programming algorithm development for more stable ReRAM and higher temperature applications

Outline

- ❖ Context and introduction
- ❖ Technology challenges and solutions
 - ◆ Bit error rate
 - ◆ Stability
- ❖ **AI opportunities**
- ❖ Conclusions

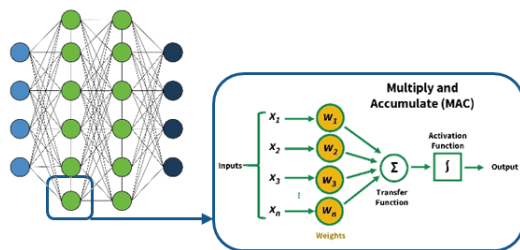


ReRAM for Neuromorphic

ReRAM perfectly fits in brain inspired systems, in a timely manner:

Embedded, Edge AI

Short
Term



ReRAM used as eNVM to store synaptic weights of the NN

Mid Term

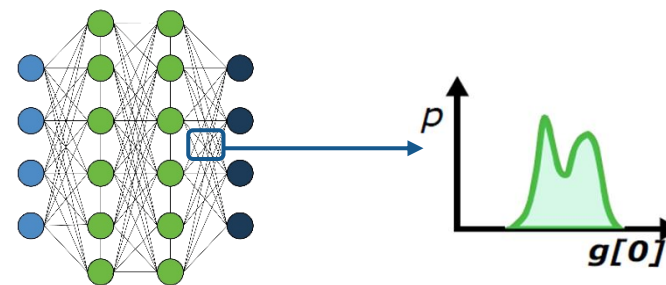
In-Memory Compute:
AI and ML



Computing is done within
the ReRAM itself

New Concepts

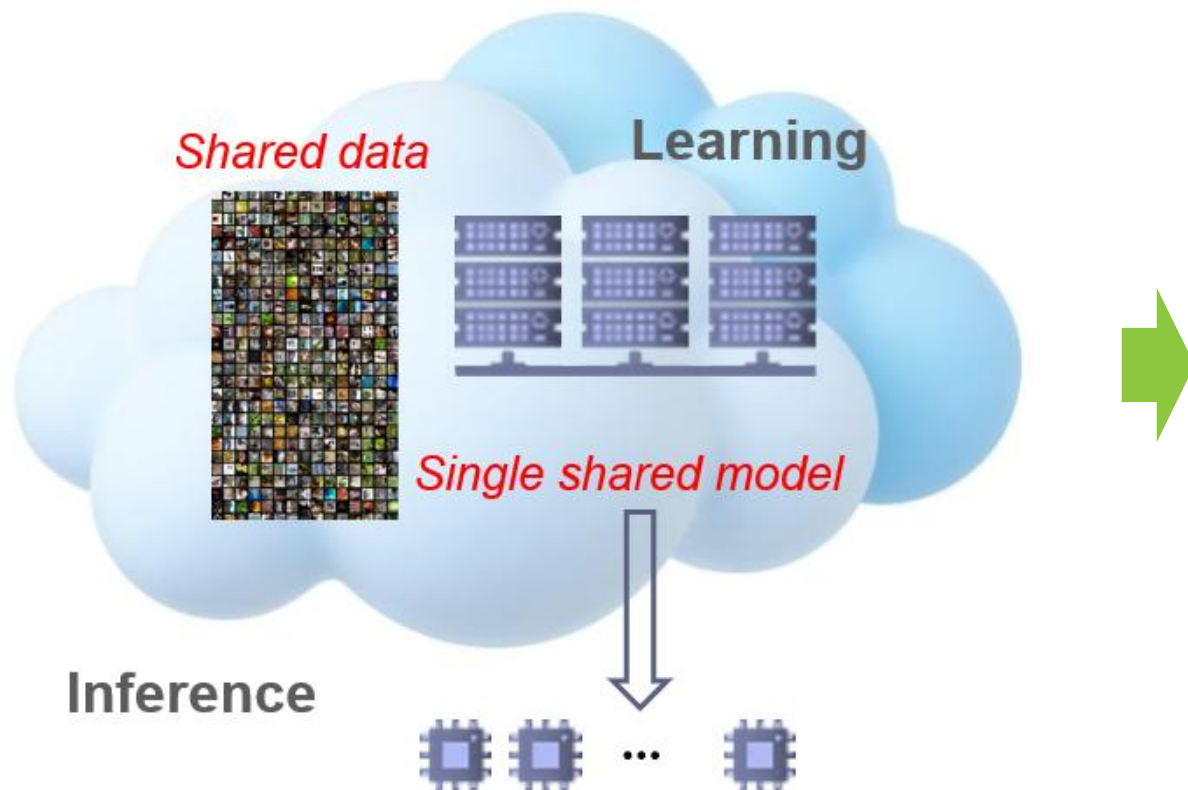
Longer
Term



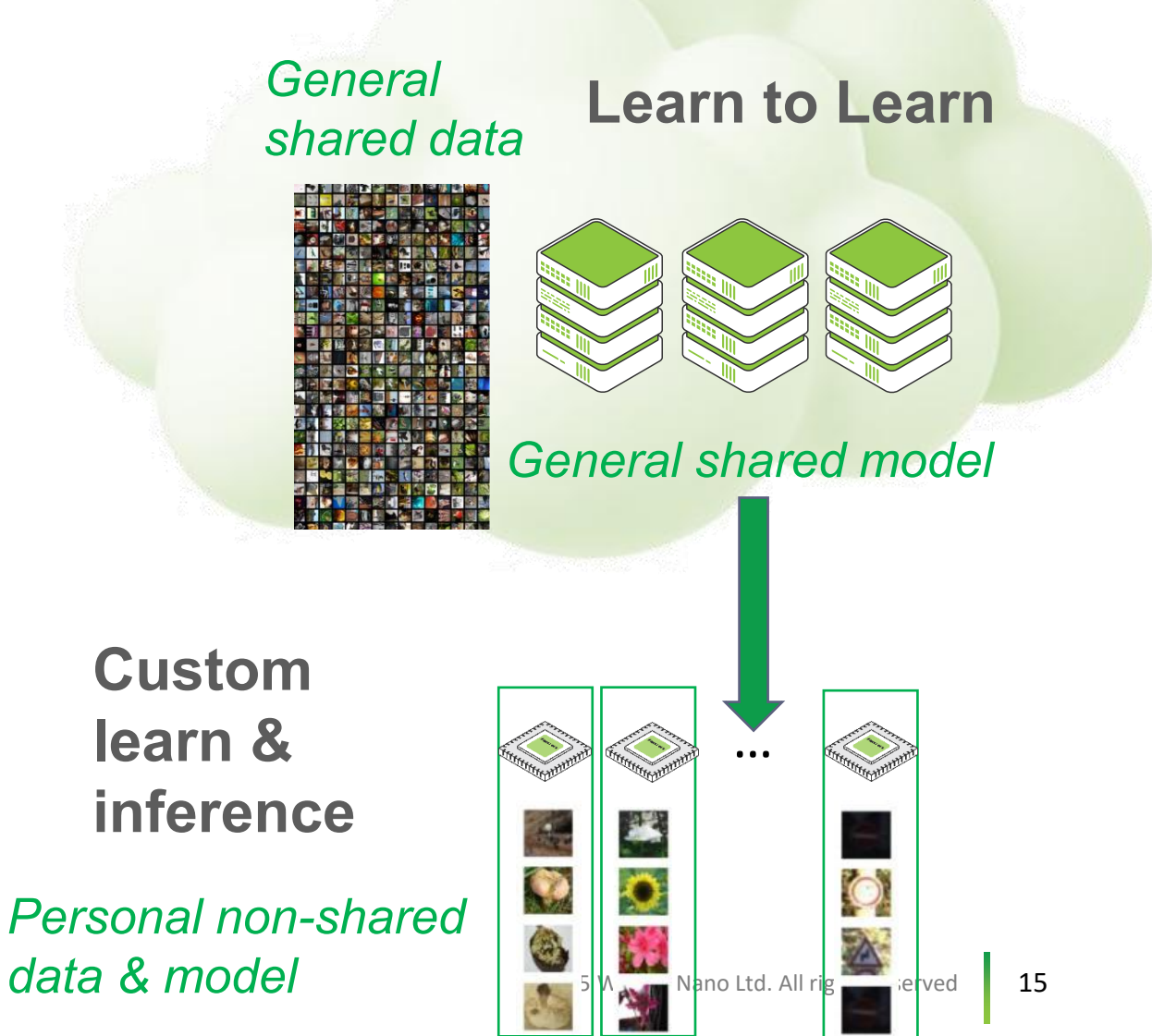
New concepts take advantage of
ReRAM specificities

Distributed Learning for Secure AI

Today: Training in the Cloud, Inference at the Edge



Tomorrow: Distributed Learning for Secure AI

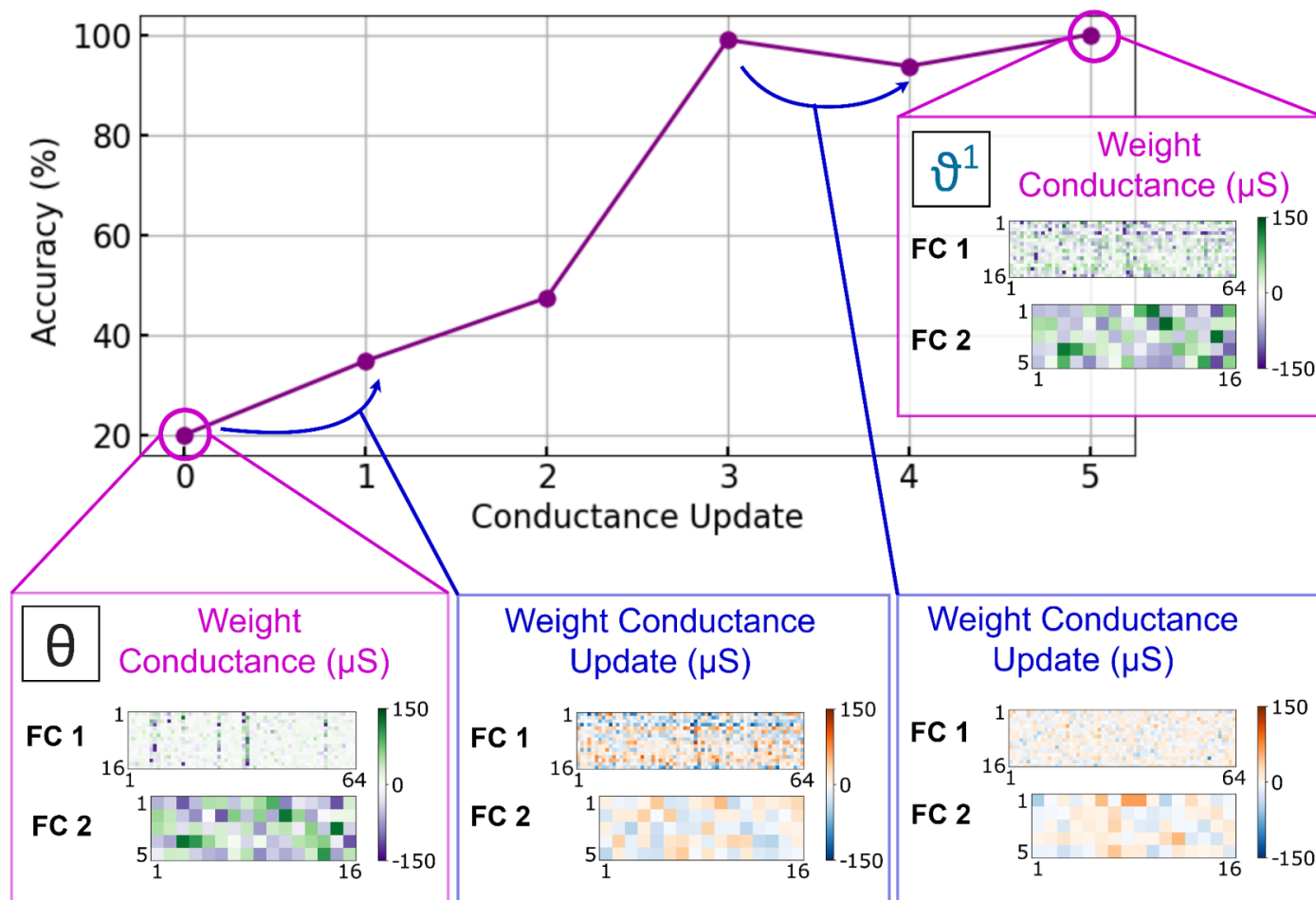


Distributed Learning for Secure AI

Model-Agnostic-Meta-Learning (MAML)

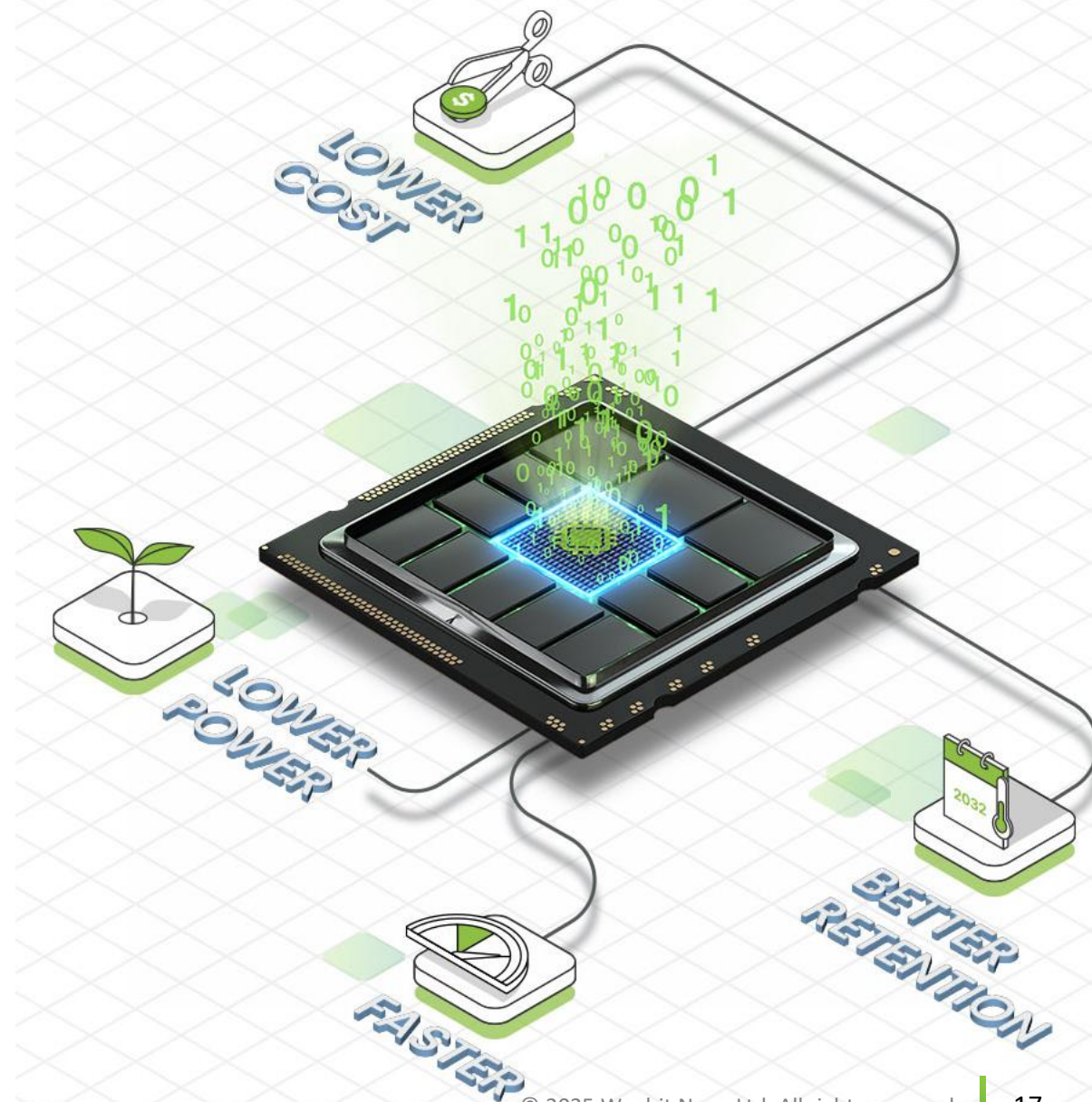
- ❖ OFF-chip:
 - ◆ MAML model trained on server (over than 50k iterations on GPU)
 - ◆ Output off-chip:
 - Optimized net parameters
- ❖ ON-chip:
 - ◆ Standalone ReRAM device able to quickly learn new tasks
 - ◆ Confidentiality assured between customers and between customer and company

Demonstration of few shot learning on ReRAM crossbar arrays



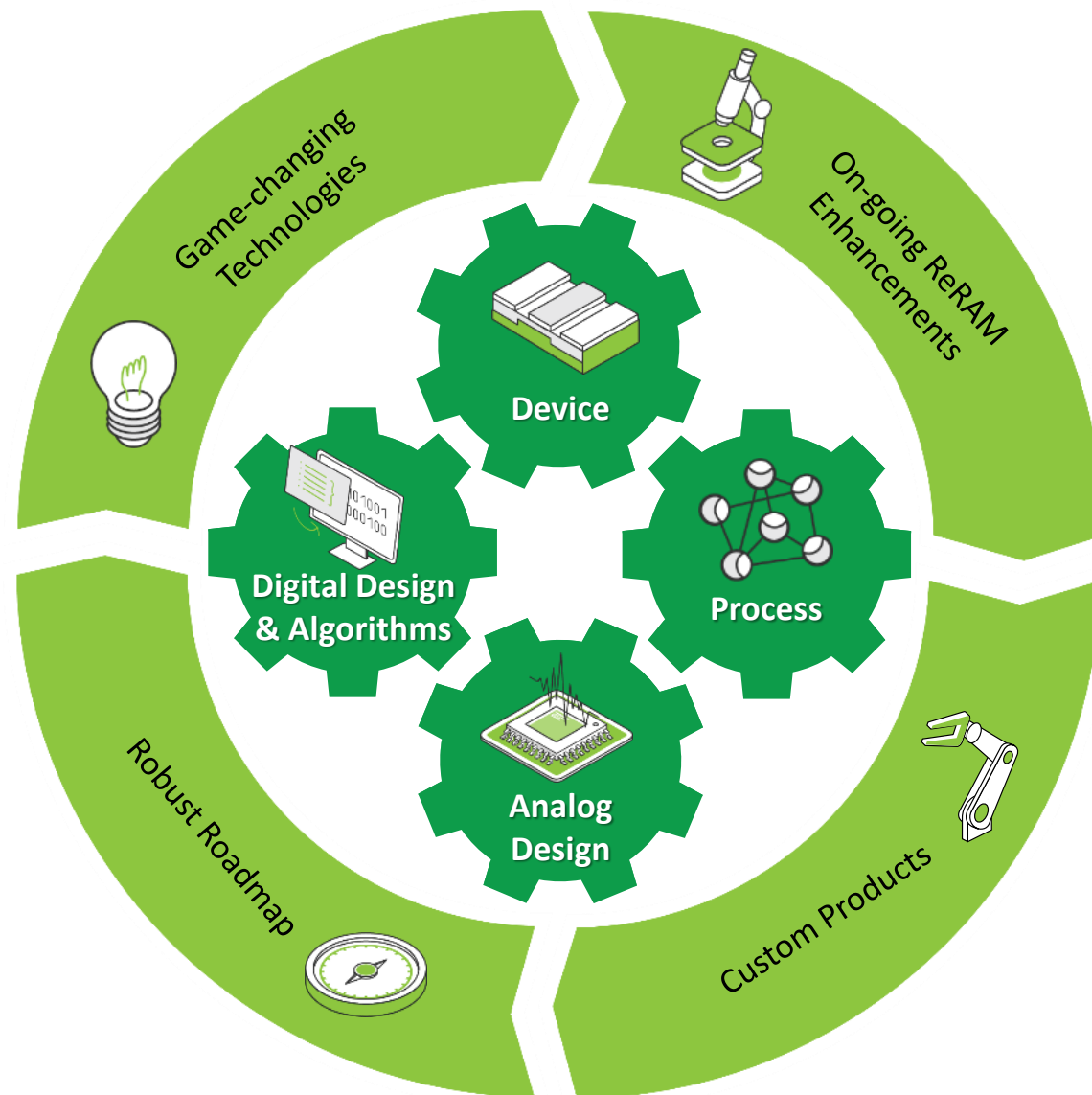
Outline

- ❖ Context and introduction
- ❖ Technology challenges and solutions
 - ◆ Bit error rate
 - ◆ Stability
- ❖ AI opportunities
- ❖ **Conclusions**



Conclusions

- ❖ ReRAM has progressed from an emerging device to a **market ready technology**, replacing embedded flash for advanced nodes
- ❖ ReRAM characteristics are perfectly suited to MCU, consumer to automotive, thanks to high speed, cost efficiency, small footprint, scalability...
- ❖ Weebit ReRAM technology is already qualified for automotive, adopted in multiple fabs and IDMs
- ❖ ReRAM will enter the **AI roadmap** step-by-step (1) as embedded memory, (2) as computing device in IMC circuits, (3) as the core device for new concepts
- ❖ ReRAM requires **continuous technology development**, involving material, process, device, test, reliability and design experts



Thank You!

www.weebit-nano.com



 **Weebitnano**
THE NEXT NVM IS HERE