Advances and Trends of RRAM technology
Gosia Jurczak
imec

September 2-4
Taipei, Taiwan
Outline

• Memory trends
• RRAM applications
• RRAM flavors: benefits and challenges
• Summary
What is RRAM?

Resistive switching material

UNIPOLAR

BIPOLAR

Low-R

High-R
Resistance modulation mechanisms

<table>
<thead>
<tr>
<th>1D</th>
<th>2D</th>
<th>3D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Filamentary</td>
<td>Interfacial</td>
<td>Bulk Transition</td>
</tr>
<tr>
<td><strong>Oxygen vacancy migration</strong></td>
<td><strong>Thermo-Chemical Fuse/antifuse</strong></td>
<td><strong>Super-Lattice TI/FE</strong></td>
</tr>
<tr>
<td><strong>Electro-Chemical ECM</strong></td>
<td><strong>Schottky or Tunnel Barrier</strong></td>
<td><strong>Phase Change PCM</strong></td>
</tr>
<tr>
<td><strong>Tunnel Magneto resistance</strong></td>
<td><strong>Electronic MIT (Mott)</strong></td>
<td></td>
</tr>
</tbody>
</table>

**1D Filamentary**
- TE: TMO: HfO2, Ta2O5, WOx
- BE: TMO: NiO2

**2D Interfacial**
- TE: Cation Source (Ag+, Cu+ or...), Metal oxide, reduced
- BE: Tunnel barrier

**3D Bulk Transition**
- TE: Poly-crystalline, amorphous
- BE: NICO, Hf/Ta, Smooth BE, Monoclinic, Hexagonal phase

**Bipolar**
- Chalcogenides alloys
- Memristor VMCO, PCMO, TiO2

**Unipolar**
- Chalcogenides alloys: GST
- STT RAM (CoFeB, MgO)
- VO2, NbO2
Possible RRAM application space

- Smart Cards
- Industrial
- Automotive
- Security
- Body Area Network
- Flexible electronics
- Domotics
- Big data
- Data Centric Computing
- Synaptic devices in neuromorphic computing
- Aerospace
- Big data
- Data Centric Computing
- Synaptic devices in neuromorphic computing
- Aerospace
- Big data
- Data Centric Computing
- Synaptic devices in neuromorphic computing
- Aerospace
- Big data
- Data Centric Computing
- Synaptic devices in neuromorphic computing
- Aerospace
- Big data
- Data Centric Computing
- Synaptic devices in neuromorphic computing
- Aerospace
Memory trends: Storage class memories

<table>
<thead>
<tr>
<th>Year</th>
<th>Technology</th>
<th>Gen 1</th>
<th>Gen 2</th>
<th>Gen 3</th>
<th>Gen 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>Micron: CBRAM IEDM2014 16Gb 27nm cell</td>
<td>32Gb 64Gb</td>
<td>128Gb 256Gb</td>
<td>512Gb</td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>IM, 3D Crosspoint July 2015: 128Gb 20nm 1S1R cell</td>
<td>PCM Gen 1</td>
<td>PCM Gen 2</td>
<td>PCM Gen 3</td>
<td>PCM Gen 4</td>
</tr>
<tr>
<td>2018</td>
<td>RRAM 1S1R, 1R</td>
<td>IM, 3D Crosspoint July 2015: 128Gb 20nm 1S1R cell</td>
<td>RRAM Gen 1</td>
<td>RRAM Gen 2</td>
<td>RRAM Gen 3</td>
</tr>
<tr>
<td>2020</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2022</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**RRAM Requirements**
- Power: low current switching <10uA
- Cost: <25% DRAM, 4F², Selector
- Capacity: >10% NAND
- Speed, endurance, retention between DRAM and NAND
- Scalability and capability of 3D vertical integration

**Notes:**
- 15nm 4F²
- Gen 2
- Gen 3
- Gen 4

**Scalability Diagram:**
- SRAM, DRAM, Latency gap, NAND Flash, Mass storage
Memory trends: Embedded NVM

### e-NVM today
- 1T NOR
- FSL, IFX, NXP, ST
- HS3P/IFX
- ESF/SST

### e-FLASH Challenges
- Incompatibilities with HK/MG
- Increased costs (+12 masks)
- High programming voltage (8-10V)
- Scalability of NOR Flash

### RRAM: Requirements
- BEOL process compatibility, scalability
- Process cost, Core CMOS voltage compatible
- Low power, low energy, cost
## Embedded NVM requirements

<table>
<thead>
<tr>
<th>Application</th>
<th>Key Attributes</th>
<th>Other requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Automotive</strong></td>
<td>▪ Array size: 1MB-16MB</td>
<td>• Needs more mature technology</td>
</tr>
<tr>
<td></td>
<td>▪ Endurance: 1E6</td>
<td>• Good understanding of failure modes, reliability mechanisms</td>
</tr>
<tr>
<td></td>
<td>▪ Retention: 20yrs @ 150C</td>
<td>• Reduced variability for zero defect</td>
</tr>
<tr>
<td></td>
<td>▪ Zero Defect capability</td>
<td>• BER low for high quality memory</td>
</tr>
<tr>
<td><strong>General Market</strong></td>
<td>▪ Array size: 16kB-4MB</td>
<td>• Need to minimize process cost</td>
</tr>
<tr>
<td></td>
<td>▪ Endurance: 1E4-1E5</td>
<td>• Voltages compatible with core transistor</td>
</tr>
<tr>
<td></td>
<td>▪ Retention: 10yrs @ 85C</td>
<td>• Compatibility with advanced HKMG CMOS thermal budget</td>
</tr>
<tr>
<td></td>
<td>▪ Retention industrial 10yrs @ 105C</td>
<td>• Compatibility with CMOS BEOL thermal budget</td>
</tr>
<tr>
<td><strong>IoT Sensor node</strong></td>
<td>▪ Array size: 16KB-512KB</td>
<td>• Cost is critical (integrated into 28-40nm CMOS node)</td>
</tr>
<tr>
<td></td>
<td>▪ Endurance: &gt;1E6</td>
<td>• Minimal write/read energy</td>
</tr>
<tr>
<td></td>
<td>▪ Retention: 10yrs @ 70C</td>
<td>• Minimal variations to minimize verify algorithms</td>
</tr>
</tbody>
</table>
## RRAM as e-Flash replacement in ULP IOT devices

<table>
<thead>
<tr>
<th></th>
<th>e-FLASH (NOR)</th>
<th>e-RRAM</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Process costs</strong></td>
<td>1T FEOL device, 12-15 additional maskset</td>
<td>1T1R (1T: I/O CMOS device, 1R:BEOL compatible)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2-3 additional masks</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Simple materials (MIM structure, no new tools/processes)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1T1R with BEOL process (Low T electronics)</td>
</tr>
<tr>
<td><strong>Cell size</strong></td>
<td>1T NOR 10-30F²</td>
<td>1T1R (voltages compatible to Core T) &lt;20F²</td>
</tr>
<tr>
<td></td>
<td>Splitgate 20-40F²</td>
<td>1T1R (voltages compatible to I/O T) 30-50F²</td>
</tr>
<tr>
<td></td>
<td>2T NOR 25-100F²</td>
<td>2T1R (Vswitch&gt;2.5V) 25-40F²</td>
</tr>
<tr>
<td></td>
<td></td>
<td>BEOL selector would enable area scaling: 1S1R scaled down to 4F2</td>
</tr>
<tr>
<td><strong>Periphery area</strong></td>
<td>Charge Pump circuits</td>
<td>Much smaller CP circuits</td>
</tr>
<tr>
<td><strong>Write energy</strong></td>
<td>V=10V, I=100uA, speed 1us E/bit ~0.5-1nJ/bit + power/energy loss on CP</td>
<td>V<del>2V, I</del>100uA-200uA, Speed 100ns E/bit ~ 40pJ/bit</td>
</tr>
<tr>
<td></td>
<td>Block erase, program with verify</td>
<td>RAM, no erase needed, no refresh</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Verification needed at lower current (variability)</td>
</tr>
<tr>
<td><strong>Read energy</strong></td>
<td>I<del>20uA, VD</del>0.5V, VG~5V, speed 35ns</td>
<td>I<del>10uA, VD</del>0.1V, VG&lt;2V, depending on SA; RW is decreasing with programming current</td>
</tr>
<tr>
<td><strong>Lifetime</strong></td>
<td>Endurance: 1e6</td>
<td>Endurance &gt;1e6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Trade-off between write energy and retention</td>
</tr>
<tr>
<td><strong>Wake up speed</strong></td>
<td>BL to be charged to 10V</td>
<td>BL to be charged to 2V</td>
</tr>
</tbody>
</table>
Which emerging memory for IoT?

Energy

Cost

STT RAM

RRAM

TE

OEL

TMO

BE

Write Energy (J)

Write Time (ns)

10^{-10}

10^{-9}

10^{-8}

10^{-7}

10^{-6}

10^{-5}

10^{-4}

10^{-3}

10^{-2}

10^{-1}

10^{0}

10^{1}

10^{2}

10^{3}

10^{4}

10^{5}

10^{6}

RRAM

CBRAM

PCM

STT-MRAM

Courtesy: P.Wong group, Stanford University
RRAM maturity

Many macros already presented
RRAM products: Panasonic MN101L

MN101L is an 8-bit microcontroller with 64kB memory
Operating T range: -40°C to +85°C
62kB of memory rated for 1e3 program cycles
2kB of memory designated for data area and rated separately for 1e5 program cycles.

SET/RESET ~2.5V, 100ns
Verify ratio ~1.2
Vread~0.5V

J.Y. Scharlotta, IIRW 2014
CBRAM

Gen1: W/GeS2/Ag/TiN

- TiN plate cap
- PolyX Ag anode
- GeS2
- W cathode

Gen2: Metal/oxide/Cu alloy

- Damacene metal plug
- 128kb-1Mb 130nm CMOS

-RESET voltage
  - 1us, 1.5V

-READ disturb (typical 0.2V)

- J. Jameson, IEDM 2013
CAS: Complementary Atom Switch for Pass Transistor in BEOL

Clermidy, DATE2014

Banno, LEAP, VLSIT2014

<table>
<thead>
<tr>
<th>Table 1 Summary of the Cu atom switch.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Buffer</td>
</tr>
<tr>
<td>Electrolyte</td>
</tr>
<tr>
<td>Inert electrode</td>
</tr>
<tr>
<td>Switching voltage</td>
</tr>
<tr>
<td>Set current (tuned by transistor)</td>
</tr>
<tr>
<td>Reset Current (peak)</td>
</tr>
<tr>
<td>Cycling endurance</td>
</tr>
<tr>
<td>OFF Disturb</td>
</tr>
<tr>
<td>Retention</td>
</tr>
</tbody>
</table>
RRAM in security applications
Reconfigurable PUF

(a) Cumulative probability vs. resistance graph showing low and high resistance states (LRS and HRS) with ±1σ and ±2σ deviations.
(b) Diagram illustrating the reconfigurable PUF architecture with address and response components.

PUF implemented on a 256kb array of RRAM in HRS, with 256b responses distribution shown in (a).

3-bit response values distribution shown in (b).

PUF based on a 256kb array of RRAM in HRS, with inter-chip Hamming Distance (HD) distribution shown in (a).

A. Chen, EDL, no2, 2015
RRAM as TRNG

Natural RTN fluctuation in RRAM

RRAM as synaptic devices in artificial neural networks

Learning process

S.Yu, Stanford, IEDM 2012

Short/long Term Memory

S.Park, Nature 2015
Radiation hardness of RRAM devices

1T1R RRAM resistive window **unaffected** by strong ionizing sources

*Stephanie L. Weeden-Wright, et al. IEEE TNS Vol 61, pp2972-2978*
Challenges at low programming current

- Reduced resistive window
  - Increased variability
  - Programming instabilities
- Degraded retention
- Reduced programming speed
- Read instabilities due to RTN
Variabilities in RRAM

Variability dependence on the programming pulse width

Variability determined by Rmin
Independent of TMO
The same variability for D2D and C2C
Programming speed

Write speed slows down with programming current due to reduced Joule heating
Retention

Z. Wei, Panasonic, IMW2012

R.Yasuhara, Panasonic, IMW2013

S. Muraoka, Panasonic, VLSI2013, 6.2

A. Calderoni, IMW 2014
Programming instabilities

OXRAM

S.Clima, IEEE EDL, no 8, 2015

CBRAM

A.Calderoni, Infos 2015
RTN

Sources of noise:
• Charge-induced (small component)
• Atomic movement-induced (depends on highest field in the filament, filament shape)
CBRAM status

Belmonte, IMW 2013

A.Calderoni, Infos 2015

RTN in CBRAM

Bake 1h 150C
Eq to 10yr at 55C
Ea 1.5eV
Non-filamentary switching RRAM

Imec VMCO: Vacancy Modulated Conductive Oxide

Unity CMOx
VMCO with 10uA programming current

Key features
- Non-filamentary switching
- Forming free (initially ON)
- Self compliant
- Self rectifying
- Analog switching behavior
- Steep ON/OFF distributions
- 3D vertical architecture compatible

B.Govoreanu, VLSI Symp on Tech, 2013
# Cell configuration for RRAM

<table>
<thead>
<tr>
<th>Cell size</th>
<th>Selector</th>
<th>Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>20F²-50F²</td>
<td>MOSFET (I/O, Core)</td>
<td>SoC e-NVM</td>
</tr>
<tr>
<td>4F²</td>
<td>Tunneling Diode, PT diode MIEC, TS, OTS, MIEC Varistor, FAST</td>
<td>SCM, 3D Cross-point array</td>
</tr>
<tr>
<td>4F²</td>
<td>Self Rectifying Cell</td>
<td>SCM, 3D Cross-point array, VRRAM</td>
</tr>
</tbody>
</table>
Selector benchmarking

Memory switching current: 1uA, 10uA

Nonlinearity $\text{NL@J_{max}}$

Current density $J_{g\text{max}} [\text{A/cm}^2]$

$\frac{1}{2} V_{op}$, $V_{op}$

$\text{NL} \frac{1}{2}$
# RRAM requirements

<table>
<thead>
<tr>
<th></th>
<th>SCM</th>
<th>SoC e-NVM MCU, IoT</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory array size</td>
<td>&gt;64Gb (128Gb)</td>
<td>Few kB – few Mb</td>
<td></td>
</tr>
<tr>
<td>Entry at technology node</td>
<td>&lt;20nm</td>
<td>&lt;40nm</td>
<td>cost</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Need 2 terminal BEOL selector</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Size of MOSFET</td>
</tr>
<tr>
<td>Cell size</td>
<td>&lt;4F²</td>
<td>&lt;40F²</td>
<td></td>
</tr>
<tr>
<td>Cell configuration</td>
<td>1S1R/1S</td>
<td>1T1R</td>
<td>Cost, array area</td>
</tr>
<tr>
<td>I&lt;sub&gt;READ&lt;/sub&gt;</td>
<td>&gt;1uA</td>
<td>&gt;1uA</td>
<td>Read speed</td>
</tr>
<tr>
<td>I&lt;sub&gt;WRITE&lt;/sub&gt;/bit</td>
<td>&lt;30uA</td>
<td>&lt;200uA</td>
<td>Energy/power, memory size</td>
</tr>
<tr>
<td>R&lt;sub&gt;off&lt;/sub&gt;/Ron</td>
<td>&gt;10 incl statistics</td>
<td>~10</td>
<td>MLC</td>
</tr>
<tr>
<td>W/R latency</td>
<td>&lt;1us/&lt;100ns</td>
<td>&lt;100ns/&lt;100ns</td>
<td></td>
</tr>
<tr>
<td>Endurance (WRITE)</td>
<td>&gt;1e9</td>
<td>&gt;1e6</td>
<td>IoT sensor node 10yr @85C</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Industrial 10yr @ 105C</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Automotive 10yr @175C</td>
</tr>
<tr>
<td>Retention</td>
<td>Days/weeks at 55C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V forming</td>
<td>Forming free (VF~Vset)</td>
<td>Forming free (VF~Vset)</td>
<td></td>
</tr>
<tr>
<td>SET/RESET Voltage</td>
<td>&lt;5V</td>
<td>&lt;1.5V</td>
<td>Size of MOSFET CP size</td>
</tr>
</tbody>
</table>
Summary

- RRAM is a promising memory option for SCM and e-NVM (IoT)
- Other applications: pass transistor in BEOL, security, neuromorphic computing, aerospace applications
- Key scientific and technological issues:
  - Current scaling (variability, instabilities, retention, speed)
  - Voltage scaling (energy, area)
  - selector