Non-Volatile Memory Technology: Directions Beyond Floating Gate Devices

Bob Gleixner and Chandra Mouli

Micron Technology
Two question to ask (for a new NVM)…

▶ What are the cost/performance characteristics required to add *significant* value in computing devices?

▶ For the leading emerging NVM’s, what materials improvements are needed to realize this value?
Computing memory hierarchy

Not so long ago...

CPU/SRAM  \rightarrow  DRAM  \rightarrow  DISK
Latency/size  \rightarrow  ns / GB  \rightarrow  ms / TB

Today...

CPU/SRAM  \rightarrow  DRAM  \rightarrow  NAND  \rightarrow  DISK
Latency/size  \rightarrow  ns / GB  \rightarrow  us / G-TB  \rightarrow  ms / TB

PC  \rightarrow  TAPE
Server

PC Server
Latency/size

Mobile
PC  \rightarrow  TAPE
Server
Where could a new memory add value?

(Assumption - you find a niche to grow)

NAND had a big “hole to fill” in 2009

<table>
<thead>
<tr>
<th>Normalized Parameters</th>
<th>DRAM</th>
<th>NAND</th>
<th>HDD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latency</td>
<td>1x</td>
<td>~500x</td>
<td>~100,000x</td>
</tr>
<tr>
<td>$/Bit</td>
<td>1x</td>
<td>~0.1x</td>
<td>~0.01x</td>
</tr>
</tbody>
</table>

A. Fazio, IEDM 2009

What do the holes look like today?

More importantly, what does this look like in 3-5 years?
Planar NAND is getting tough to scale...

Threshold Voltage
G. Naso, ISSCC 2013 – 128Gb 20nm NAND

Prall and Parat, IEDM 2010

3 bit/cell

Sigma

RTS Vt Range - Scaling Trend
Prall and Parat, IEDM 2010
…but continues to find solutions

Planar floating gate cell

Goda and Parat, IEDM 2012

16nm NAND announced

3D NAND

Vertical 3D NAND

Horizontal 3D NAND

Goda, VLSI-TSA 2013

Samsung Starts Mass Producing Industry’s First 3D Vertical NAND Flash

Saoul, Korea on Aug. 6, 2013

www.samsung.com

Micron customers won’t have to wait long for 3D flash memory

CEO Mark Durcan tells CNET that the company will start providing samples of the advanced memory technology to customers in the first quarter of 2013

www.cnet.com
While DRAM continues to scale

Seon Yong Cha, IEDM short course, 2011
Main vector in DRAM/NAND: Cost reduction

Today:

Latency

Bit/$

3 Years?:

Latency

Bit/$
Where can an emerging memory fit in?

- **NVDRAM** (Non-Volatile DRAM)
- **SCM** (Storage class memory)
- **NAND REPLACEMENT** (STORAGE)
Consider the “top 4” emerging memories

<table>
<thead>
<tr>
<th>Technology</th>
<th>Company</th>
<th>Niche</th>
</tr>
</thead>
<tbody>
<tr>
<td>RRAM</td>
<td>Adesto</td>
<td>EEPROM Replacement</td>
</tr>
<tr>
<td>MRAM</td>
<td>Everspin</td>
<td>Ultra high Reliability BBSRAM replacement</td>
</tr>
<tr>
<td>STRAM</td>
<td>Crocus</td>
<td>Embedded Security</td>
</tr>
<tr>
<td>FERAM</td>
<td>TI</td>
<td>Embedded</td>
</tr>
<tr>
<td>FERAM</td>
<td>Ramtron</td>
<td>Low Density – Power Meters</td>
</tr>
<tr>
<td>PCM</td>
<td>Samsung Micron</td>
<td>NOR replacement- Wireless applications</td>
</tr>
</tbody>
</table>

Prall, IMW 2012

**FIGURE 15:** Memory capacity of emerging nonvolatile memories.

IEEE Solid-State Circuits Mag. Spring 2013
Memory “scorecard”

- **Cost / bit**
  - Small bit cell
  - Low programming energy
  - Multi-level or multi-layer

- **Latency**
  - Fast read (large signal + memory window)
  - Fast / deterministic writes
  - Write endurance

<table>
<thead>
<tr>
<th></th>
<th>Planar NAND</th>
<th>DRAM</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cost</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell size</td>
<td>4F^2 16nm</td>
<td>6F^2 2xnm</td>
</tr>
<tr>
<td>Programming energy</td>
<td>Low</td>
<td>(Refresh)</td>
</tr>
<tr>
<td>Multi-level / layer</td>
<td>1-3 BPC</td>
<td>None</td>
</tr>
<tr>
<td><strong>Latency</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast read</td>
<td>10uS (page)</td>
<td>~10ns</td>
</tr>
<tr>
<td>Fast write</td>
<td>1mS (page)</td>
<td>~10ns</td>
</tr>
<tr>
<td>Endurance</td>
<td>1-100k</td>
<td>&gt;1e15</td>
</tr>
</tbody>
</table>
Phase Change Memory (PCM)

- **Storage mechanism**
  - Amorphous/poly-crystal phase of chalcogenide alloy (Ge$_2$Sb$_2$Te$_5$ – GST)

- **Writing mechanism**
  - Current-induced Joule heating

- **Sensing signal**
  - Resistance change of the GST

- **Cell structure**
  - 1 transistor, 1 resistor (1T/1R)
History of PCM Silicon Development

- **PCM cell**
  - S. Lai and T. Lowrey, IEDM 2001
  - F. Pellizzer et al., VLSI 2004
  - F. Pellizzer et al., VLSI 2006
  - G. Servalli, IEDM 2009

- **PCM array & chip**
  - M. Gill et al., ISSCC 2002
  - G. Casagrande et al., VLSI 2004
  - Bedeschi et al., ISSCC 2008
  - C. Villa et al., ISSCC 2010

Timeline:
- 2001
- 2003
- 2005
- 2007
- 2009
- 2011

Phases:
- Concept
- Demonstration
- Technology Validation
- Product Manufacturing

©2013 Micron Technology, Inc.
# PCM Scorecard

<table>
<thead>
<tr>
<th></th>
<th>Planar NAND</th>
<th>DRAM</th>
<th>PCM</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cost Overall</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell size</td>
<td>4F² 16nm</td>
<td>6F² 2x nm</td>
<td>4F² 20nm</td>
</tr>
<tr>
<td>Programming energy</td>
<td>Low</td>
<td>(Refresh)</td>
<td>~High I</td>
</tr>
<tr>
<td>Multi-level / layer</td>
<td>1-3 BPC</td>
<td>None</td>
<td>Level/Layer</td>
</tr>
<tr>
<td><strong>Latency Overall</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast read</td>
<td>10μS (page)</td>
<td>~10ns</td>
<td>&lt;100ns</td>
</tr>
<tr>
<td>Fast write</td>
<td>1mS (page)</td>
<td>~10ns</td>
<td>&gt;100ns</td>
</tr>
<tr>
<td>Endurance</td>
<td>1-100k</td>
<td>&gt;1e15</td>
<td>1e8</td>
</tr>
</tbody>
</table>
PCM Cell Challenges

- High programming current
- Resistance drift (MLC)
- Programming speed to SET/crystallized state
- Thermal disturb

Sadeghipour, ITERM 2006
Kang, IEDM 2011
Ielmini TED 2007
Choi, ISSCC 2012
Russo TED 2008
<table>
<thead>
<tr>
<th>RRAM</th>
<th>Material Systems</th>
<th>Switching mechanism</th>
<th>Comments</th>
</tr>
</thead>
</table>
| RRAM-M (CBRAM) | Cu/GdO\textsubscript{x}  
Ag/SiO\textsubscript{2}  
Ag/GeS | +V  
Cu, Ag etc.,  
W, TiN etc.,  
Metal Filament in Electrolyte | Ion/Ioff $\sim$ 1000  
K.A. Aratani et al., IEDM 2008 |
| Bipolar RRAM-Ox | Ti/HfO\textsubscript{x}, Ti/ZrO\textsubscript{x} | +V  
Ti, Hf  
TiN | Oxygen vacancy filament in dielectric TiO\textsubscript{x}, ZrO\textsubscript{x}, TaO\textsubscript{x} etc.,  
H.Y. Lee et al., IEDM 2008 |
| RRAM-MVO   | Mixed Valence Oxides  
SrTiO\textsubscript{x}, PCMO etc., | Uniform oxygen movement | Ion/Ioff $\sim$ 10  
M. Meyer et al., NVMTS 2008 |
### RRAM Scorecard

<table>
<thead>
<tr>
<th></th>
<th>Planar NAND</th>
<th>DRAM</th>
<th>RRAM</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cost Overall</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell size</td>
<td>4F&lt;sup&gt;2&lt;/sup&gt; 16nm</td>
<td>6F&lt;sup&gt;2&lt;/sup&gt; 2x nm</td>
<td>4F&lt;sup&gt;2&lt;/sup&gt; 20nm</td>
</tr>
<tr>
<td>Programming energy</td>
<td>Low</td>
<td>(Refresh)</td>
<td>Low - Med</td>
</tr>
<tr>
<td>Multi-level / layer</td>
<td>1-3 BPC</td>
<td>None</td>
<td>Level/Layer</td>
</tr>
<tr>
<td><strong>Latency Overall</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast read</td>
<td>10uS (page)</td>
<td>~10ns</td>
<td>&lt;100ns</td>
</tr>
<tr>
<td>Fast write</td>
<td>1mS (page)</td>
<td>~10ns</td>
<td>~10nS</td>
</tr>
<tr>
<td>Endurance</td>
<td>1-100k</td>
<td>&gt;1e15</td>
<td>1e8 ??</td>
</tr>
</tbody>
</table>
RRAM Cell Challenges

- **Window for MLC**
  - Write current scalability?
  - Noise

  - Write
  - Read

Chen IEDM 2012 – HfOx

Aratani IEDM 2007 -- CBRAM

Prall IMW 2012 – HfOx
MRAM

- Fast switching could enable DRAM replacement

  Fast 35 ns read/write cycle
  SRAM compatible timing
  Unlimited read & write endurance
  Data always non-volatile for >20 years at temperature
  RoHS-compliant small footprint BGA and TSOP2 package
  AEC-Q100 Grade 1 option in TSOP2 package.

- But field / toggle MRAM cell too large → STT-MRAM

  Rizzo, IEEE Trans Mag 2013
  Sun, IEEE DRC 2011

MR4A16B

1M x 16 MRAM

FEATURES

- +3.3 Volt power supply
- Fast 35 ns read/write cycle
- SRAM compatible timing
- Unlimited read & write endurance
- Data always non-volatile for >20 years at temperature
- RoHS-compliant small footprint BGA and TSOP2 package
- AEC-Q100 Grade 1 option in TSOP2 package.

PMA Stack

20nm Ru
0.25Co/0.8Pd * 14
0.9 Ru
0.3 Co
0.25Co/0.8Pd * 4
0.8 Pt
0.25Co
0.3 Ta
0.8 CoFeB
0.5 Fe
0.9 MgO
0.8 CoFeB
2 Ta
5 RuCoFe
Substrate


©2013 Micron Technology, Inc.
## STT-MRAM Scorecard

<table>
<thead>
<tr>
<th></th>
<th>Planar NAND</th>
<th>DRAM</th>
<th>STT-MRAM</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cost Overall</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell size</td>
<td>4F² 16nm</td>
<td>6F² 2x nm</td>
<td>6F² 20nm</td>
</tr>
<tr>
<td>Programming energy</td>
<td>Low (Refresh)</td>
<td>Med I</td>
<td></td>
</tr>
<tr>
<td>Multi-level / layer</td>
<td>1-3 BPC</td>
<td>None</td>
<td>Layer</td>
</tr>
<tr>
<td><strong>Latency Overall</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast read</td>
<td>10µS (page)</td>
<td>~10ns</td>
<td>&lt;30ns</td>
</tr>
<tr>
<td>Fast write</td>
<td>1mS (page)</td>
<td>~10ns</td>
<td>~10ns</td>
</tr>
<tr>
<td>Endurance</td>
<td>1-100k</td>
<td>&gt;1e15</td>
<td>&gt;1e15</td>
</tr>
</tbody>
</table>
STT-MRAM Cell Challenges

▶ Writing current vs. stability

- Main Challenge:
  - Reduce $I_{set}$ ($I_s$)
  - Maintain Distributions
  - Maintain Energy Barrier

- Cumulative distribution

Slaughter, IEDM 2012

- Thermal stability factor ($\alpha$) vs. MTJ dimension (nm)

Park, VLSI 2012

▶ Tunnel junction reliability

Gallagher, ICSICT 2012

▶ Operating window

- CoFeB/MgO/CoFeB MTJ
- perpendicular MgO-MTJ
- amorphous Al:O-MTJ

Ando, NVMTS 2011

- Log(N)
- Applied Voltage

Slaughter, IEDM 2012
FERAM

- Stores “charge” in a polarizable material
  - Typically PZT

- As with field/toggle MRAM, performance is similar to DRAM

1T1C Memory Cell Schematic

Wouters, CREMSI Workshop 2004

FM23MLD16
8Mbit F-RA Memory

Features
- 8Mbit Ferroelectric Nonvolatile RAM
  - Organized as 512Kx16
  - Configurable as 1Mx8 Using UB/LB
  - High Endurance 100 Trillion (10¹²) Read/Writes
  - NoDelay™ Writes
  - Page Mode Operation to 31MHz
  - Advanced High-Reliability Ferroelectric Process

SRAM Compatible
- JEDEC 512Kx16 SRAM Pinout
- 60 ns Access Time, 115 ns Cycle Time

Advanced Features
- Low Vth Monitor Protects Memory against Inadvertent Writes

Superior to Battery-backed SRAM Modules
- No Battery Concerns
- Monolithic Reliability
- True Surface Mount Solution, No Rework Steps
- Superior for Moisture, Shock, and Vibration

Low Power Operation
- 2.7V – 3.6V Power Supply
- 14 mA Active Current

Industry Standard Configuration
- Industrial Temperature -40°C to +85°C
- 45-pin “Green” RoHS FBGA package

http://www.cypress.com/
# FERAM Scorecard

<table>
<thead>
<tr>
<th></th>
<th>Planar NAND</th>
<th>DRAM</th>
<th>FERAM</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cost Overall</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell size</td>
<td>4F² 16nm</td>
<td>6F² 2x nm</td>
<td>&gt;100nm</td>
</tr>
<tr>
<td>Programming energy</td>
<td>Low</td>
<td>(Refresh)</td>
<td>Low</td>
</tr>
<tr>
<td>Multi-level / layer</td>
<td>1-3 BPC</td>
<td>None</td>
<td>Layer?</td>
</tr>
<tr>
<td><strong>Latency Overall</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast read</td>
<td>10uS (page)</td>
<td>~10ns</td>
<td>~10ns</td>
</tr>
<tr>
<td>Fast write</td>
<td>1mS (page)</td>
<td>~10ns</td>
<td>~10ns</td>
</tr>
<tr>
<td>Endurance</td>
<td>1-100k</td>
<td>&gt;1e15</td>
<td>&gt;1e14</td>
</tr>
</tbody>
</table>
FERAM Cell Challenges

- Need to reduce cells size - limited by polarization / bitline cap ratio

- Considering cell / array architecture solutions

<table>
<thead>
<tr>
<th>Design Rule</th>
<th>0.5μm</th>
<th>0.25μm</th>
<th>0.2μm</th>
<th>130nm</th>
<th>128nm+DualOx</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capacity</td>
<td>16Kb</td>
<td>8Mb</td>
<td>32Mb</td>
<td>64Mb</td>
<td>128Mb</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell Capacitor</td>
<td>Spatter PZT</td>
<td>Spatter PZT</td>
<td>Spatter PZT</td>
<td>MOCCVD PZT</td>
<td>MOCCVD PZT</td>
</tr>
<tr>
<td>Cap. Size</td>
<td>1.24μm²</td>
<td>0.855μm²</td>
<td>0.49μm²</td>
<td>0.134μm²</td>
<td>0.038μm²</td>
</tr>
<tr>
<td>Cell Size</td>
<td>15.8μm²</td>
<td>5.2μm²</td>
<td>1.87μm²</td>
<td>0.719μm²</td>
<td>0.32μm²</td>
</tr>
<tr>
<td>Process</td>
<td>2Metal</td>
<td>2Metal</td>
<td>3Metal</td>
<td>3Metal</td>
<td>3Metal</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4Metal Dual Ox</td>
</tr>
<tr>
<td>Chip Size</td>
<td>1.9mm²</td>
<td>2.0mm²</td>
<td>2.8mm²</td>
<td>2.87mm²</td>
<td>2.87mm²</td>
</tr>
<tr>
<td>R/W Cycle</td>
<td>30ns</td>
<td>17ns</td>
<td>7.5ns</td>
<td>60ns</td>
<td>70ns</td>
</tr>
<tr>
<td>Page/Burst</td>
<td>80ns</td>
<td>30ns</td>
<td>7.5ns</td>
<td>10ns</td>
<td>1.25ns</td>
</tr>
<tr>
<td>R/W Bandwidth</td>
<td>125MB/s</td>
<td>50MB/s</td>
<td>80MB/s</td>
<td>200MB/s</td>
<td>1.6GB/s</td>
</tr>
<tr>
<td>Vddl</td>
<td>3.3V</td>
<td>3.3V</td>
<td>3.3V/2.5V</td>
<td>3.3V/2.5V</td>
<td>1.8V</td>
</tr>
</tbody>
</table>

Takashima NVMTS 2011

Shiga JSSC 2010
Multi-Layer Memory

- Crossbar memory very attractive
  - “Simple” structure and minimum cell size (4F²)
  - Suitable for 3D stacking → cell size (4/n)F²
  - Array over circuitry → better array efficiency

- The basic cell architecture requires a selector structure to be integrated in the BEOL
  - Parasitic paths exist through neighboring cells
  - Programming (and also reading) can perturb the array
A Wide Range of Material Choices

For the selector structure several devices have been proposed so far, none have been “proven”

Selector device options

- Homojunctions → polySi p/n junctions
- Heterojunctions → p-CuO/n-InZnO
- Schottky diode → Ag/n-ZnO
- Chalcogenide Ovonic Threshold Switching (OTS) materials
- Mixed Ionic Electronic Conduction (MIEC) materials
Cross-Point Switch Requirements

- Very high forward bias current
  - Greater than the switching current
- Low leakage through unselected cells
  - Steep IV characteristic
- Bipolar operation may be required
  - Necessary for bipolar RRAM

<table>
<thead>
<tr>
<th>Cell Type</th>
<th>~20nm Cell Current</th>
<th>Current density</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCM</td>
<td>100μA</td>
<td>10 MA/cm²</td>
</tr>
<tr>
<td>RRAM</td>
<td>10μA</td>
<td>1 MA/cm²</td>
</tr>
</tbody>
</table>

Example for read

Array size < $I_{LRS}(V) / I_{LRS}(V/2)$
# Summary Scorecard

<table>
<thead>
<tr>
<th></th>
<th>Planar NAND</th>
<th>DRAM</th>
<th>PCM</th>
<th>RRAM</th>
<th>STT-MRAM</th>
<th>FERAM</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cost Overall</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell size</td>
<td>4F² 16nm</td>
<td>6F² 2x nm</td>
<td>4F² 20nm</td>
<td>4F² 20nm</td>
<td>6F² 20nm</td>
<td>&gt;100nm</td>
</tr>
<tr>
<td>Programming energy</td>
<td>Low (Refresh)</td>
<td>~High I</td>
<td>Low - Med</td>
<td>Med I</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Multi-level / layer</td>
<td>1-3 BPC</td>
<td>None</td>
<td>Level/ Layer</td>
<td>Level/ Layer</td>
<td>Layer</td>
<td>Layer?</td>
</tr>
<tr>
<td><strong>Latency Overall</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast read</td>
<td>10uS (page)</td>
<td>~10ns</td>
<td>&lt;100ns</td>
<td>&lt;100ns</td>
<td>&lt;30ns</td>
<td>~10ns</td>
</tr>
<tr>
<td>Fast write</td>
<td>1mS (page)</td>
<td>~10ns</td>
<td>&gt;100ns</td>
<td>~10ns</td>
<td>~10ns</td>
<td>~10ns</td>
</tr>
<tr>
<td>Endurance</td>
<td>1-100k</td>
<td>&gt;1e15</td>
<td>1e8</td>
<td>1e8 ??</td>
<td>&gt;1e15</td>
<td>&gt;1e14</td>
</tr>
</tbody>
</table>
Where can an emerging memory fit in?

- **NVDRAM** (Non-Volatile DRAM)
- **SCM** (Storage class memory)
- **NAND REPLACEMENT** (STORAGE)
Summary

- Many new NVM technologies are attempting to gain a foothold in the computing devices
- While they have cost or performance advantages over NAND or DRAM, none beat the performance of DRAM or the cost of NAND
- The most direct fit would be a NAND/DRAM “blend” / SCM, but the system implementation for this type of memory is not yet established
- We need a few more advances in materials to make these a reality!

Acknowledgments: Thanks to the Micron R&D team and the work of the authors presented here.