Emerging Memory Update 2019: What a Difference a Year Makes!

Jim Handy, Objective Analysis
Tom Coughlin, Coughlin Associates
What A Difference A Year Makes!
Outline

- Emerging Memory Update by Type
- Emerging Memory Support Update
- Hurdles & Outlook
Outline

- Emerging Memory Update by Type
- Emerging Memory Support Update
- Hurdles & Outlook
Emerging Memory Types

- PCM/XPoint
- MRAM
- ReRAM
- FRAM
- Others
PCM/3D XPoint – “Optane”
3D XPoint Must Be Priced Below DRAM Otherwise People will Just Buy DRAM
Intel Incurring Significant XPoint Losses

<table>
<thead>
<tr>
<th>Year</th>
<th>Net Profit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1Q16</td>
<td>-60%</td>
</tr>
<tr>
<td>2Q16</td>
<td>-40%</td>
</tr>
<tr>
<td>3Q16</td>
<td>-20%</td>
</tr>
<tr>
<td>4Q16</td>
<td>0%</td>
</tr>
<tr>
<td>1Q17</td>
<td>20%</td>
</tr>
<tr>
<td>2Q17</td>
<td>40%</td>
</tr>
<tr>
<td>3Q17</td>
<td>60%</td>
</tr>
<tr>
<td>4Q17</td>
<td>0%</td>
</tr>
<tr>
<td>1Q18</td>
<td>20%</td>
</tr>
<tr>
<td>2Q18</td>
<td>40%</td>
</tr>
<tr>
<td>3Q18</td>
<td>60%</td>
</tr>
<tr>
<td>4Q18</td>
<td>0%</td>
</tr>
<tr>
<td>1Q19</td>
<td>20%</td>
</tr>
<tr>
<td>2Q19</td>
<td>40%</td>
</tr>
</tbody>
</table>

Legend:
- Intel
- Samsung
- Micron
- SK hynix
- Toshiba
- WDC
Meanwhile, DRAM Prices are Collapsing

-59%
PCM Status

- Optane SSDs gaining modest acceptance
  - NAND makers countering with fast SLC SSDs
- Optane DIMMs key selling point for next-generation server CPUs
3D XPoint Report

- 2019 Update from Objective Analysis
- The Why, How, and When of 3D XPoint Memory
  - Why Intel wants it
  - How it fits into the memory hierarchy
    - Impact on DRAM
  - When will it sell in volume
- Detailed Forecasts

https://Objective-Analysis.com/reports/#XPoint
Magnetic RAM: MRAM
Three MRAM Types: Toggle, STT, & SOT

**Toggle Mode (legacy)**

- **“Reset”**
  - "Free" magnetic layer
  - "Fixed" magnetic layer
  - Magnets aligned = Low resistance: “0”

- **“Set”**
  - Magnets unaligned = High resistance. “1”

**Spin Transfer Torque (STT, ramping)**

- **Low Resistance or “0”**
  - Parallel Magnetic Polarization
  - Free Layer
  - Tunnel Barrier
  - Reference Layer

- **High Resistance or “1”**
  - Anti-Parallel Magnetic Polarization

2019 Storage Developer Conference. © Objective Analysis & Coughlin Associates. All Rights Reserved.
Spin-Orbit Torque MRAM (future)

- Higher reliability in-plane current
- SOT switching faster than STT

SOT MRAM could achieve full SRAM speeds
MRAM Status

- MRAM cache in IBM SSDs
- Spin Memory (Spin Torque Technologies) launches:
  - Precessional Spin Current
  - Endurance Engine Technologies
- All major foundries sampling MRAM, two shipping
- Everspin still sole supplier of stand-alone MRAMs
  - Over 123 million units shipped
  - Avalanche is sampling
- Today’s markets: Space, high-uptime systems, caches and buffers
Embedded MRAM (1)

- MRAM will shrink past SRAM & flash
- Tuned to application
  - Retention
  - Endurance
  - Capacity
Embedded MRAM (2)

- 1T/2T MRAM smaller than 6T SRAM
- NOR flash scaling limit ≤15nm
- MRAM could replace embedded SRAM & NOR
  - Lower power
  - Lower cost
  - Higher density

Source: Kevin Moraes, AMAT, 2019
MRAM Foundry Developments

- All major semiconductor foundries involved
  - Samsung, TSMC, Global Foundries, UMC…
- Many plan to cost reduce
  - Moving from BEOL to front end
- MRAM needs new tools
  - Drives more capital spending
Resistive RAM: ReRAM
Many Flavors of ReRAM

- CMOx
- CBRAM
- PCM
- Memristor
- Carbon nanotubes

All use the value of a resistance to represent a “1” or “0”
Two ReRAM Types in the Lead

- SiO$_2$-based
  - Crossbar, Weebit Nano
    - Positioned as the memory for neural networks
    - Crossbar production at SMIC
  - Not yet in volume
- Metal filament
  - Adesto in volume production
Ferroelectrics: FRAM
Ramtron (Now Cypress)
  - Sole supplier of stand-alone FRAM
    - PZT – Lead Zirconium Titanate

Other renditions:
  - Thinfilm, organic FRAMs
  - Symetrix IP provider

New HfO₂ approach from NamLab, Dresden
  - Uses well-understood materials (Hafnium Oxide)

Today’s markets:
  - RFID, other low write current applications
All New Memories Share Some Attributes

- Small single-element cell
  - Supports small/inexpensive die and 3D stacking
  - Promises to scale past DRAM & NAND flash
- Write in place
  - No “Block Erase”
  - More symmetrical read/write speeds
- Nonvolatile/Persistent
  - These can all be used as Persistent Memory: “PM”
# Memory Attributes By Technology

<table>
<thead>
<tr>
<th>Units</th>
<th>SRAM</th>
<th>DRAM</th>
<th>NOR Flash</th>
<th>NAND Flash</th>
<th>Toggle</th>
<th>ST-MRAM</th>
<th>FRAM</th>
<th>PCM</th>
<th>RRAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Byte Read Time</td>
<td>ns</td>
<td>2</td>
<td>10</td>
<td>25</td>
<td>10,000</td>
<td>35</td>
<td>&lt;10</td>
<td>55</td>
<td>16</td>
</tr>
<tr>
<td>Byte Write Time</td>
<td>ns</td>
<td>2</td>
<td>10</td>
<td>5,000,000</td>
<td>200,000</td>
<td>35</td>
<td>&lt;10</td>
<td>55</td>
<td>30</td>
</tr>
<tr>
<td>Standby Current</td>
<td>mA</td>
<td>&lt;1</td>
<td>45</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Read Current</td>
<td>mA</td>
<td>20</td>
<td>220</td>
<td>20</td>
<td>25</td>
<td>30</td>
<td>15</td>
<td>&lt;10</td>
<td>15</td>
</tr>
<tr>
<td>Write Current</td>
<td>mA</td>
<td>20</td>
<td>240</td>
<td>50</td>
<td>25</td>
<td>30</td>
<td>15</td>
<td>&lt;10</td>
<td>20</td>
</tr>
<tr>
<td>Endurance</td>
<td>P/E Cycles</td>
<td>Infinite</td>
<td>Infinite</td>
<td>$10^5$</td>
<td>$10^4$</td>
<td>$10^{11}$</td>
<td>$10^{13}$</td>
<td>$10^{14}$</td>
<td>$10^6$</td>
</tr>
<tr>
<td>Retention</td>
<td>Yrs @ 55°C</td>
<td>0</td>
<td>10^{-9}</td>
<td>&gt;10</td>
<td>&gt;10</td>
<td>&gt;20</td>
<td>&gt;20</td>
<td>&gt;10</td>
<td>&gt;10</td>
</tr>
<tr>
<td>Scaling Limit</td>
<td>nm</td>
<td>5</td>
<td>10?</td>
<td>45?</td>
<td>14</td>
<td>65</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Cell Size</td>
<td>f²</td>
<td>50</td>
<td>6-8</td>
<td>6-8</td>
<td>4 effective with MLC</td>
<td>35-40</td>
<td>8-9</td>
<td>8-20</td>
<td>4</td>
</tr>
<tr>
<td>Select Device</td>
<td>N/A</td>
<td>Transistor</td>
<td>Transistor</td>
<td>Transistor</td>
<td>Transistor</td>
<td>Transistor</td>
<td>Transistor</td>
<td>Transistor</td>
<td>Diode</td>
</tr>
<tr>
<td>MLC Capability</td>
<td>Bits/Cell</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>SEU Immune</td>
<td>N/A</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>SEL Immune</td>
<td>N/A</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>TID</td>
<td>krad (Si)</td>
<td>&lt;100</td>
<td>&lt;100</td>
<td>&lt;100</td>
<td>&lt;100</td>
<td>&gt;1,000</td>
<td>&gt;1,000</td>
<td>&lt;100</td>
<td>&gt;1,000</td>
</tr>
</tbody>
</table>
Outline

- Emerging Memory Update by Type
- Emerging Memory Support Update
- Hurdles & Outlook
Support Requirements

- Hardware advancements (JEDEC, Others)
  - Supporting early development
  - Ongoing requirements
- Software support (SNIA)
  - O/S support
  - Application program support
- EDA support
Hardware: Early Development

- Early groundwork has been helpful
  - NVDIMM-N
    - DRAM with flash backup
  - BIOS changes
    - Boot without reloading memory
  - New power fail signal brought to DIMM
- 3D XPoint driving changes
Design Tools MRAM Options

- Embedded MRAM needs EDA support
  - Synopsis support announced for 2Q19
  - Cadence to support MRAM in DDR4 controllers

Source: Synopsis, October 2018
NVDIMM Report

- Objective Analysis
- Explains the NVDIMM markets
  - NVDIMM-N
  - NVDIMM-P
- Vendor profiles
- Support requirements
- Market forecast

https://Objective-Analysis.com/reports/#NVDIMM
Ongoing Hardware Requirements

- Nonuniform Memory Architecture: “NUMA”
- MMU Redesign
- Faster context switches needed
  - Use polling for now
- Updated DDR4 bus
  - Support for non-deterministic access times
Software: Operating System Support

- SNIA’s Persistent Memory Programming Model
  - [https://www.SNIA.org/PM](https://www.SNIA.org/PM)
Software: Application Program Support

- PM is useless if its advantage is untapped
  - Persistence is unknown in most software
- This change will take some time
  - Closed systems can use it now
    - Hyperscale Data Centers, SANs
  - Open systems will evolve
Outline

- Emerging Memory Update by Type
- Emerging Memory Support Update
- Hurdles & Outlook
The Vision: Replace Existing Technologies

<table>
<thead>
<tr>
<th>Process Geometry</th>
<th>Relative Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>500nm</td>
<td>$1,000.00</td>
</tr>
<tr>
<td>250nm</td>
<td>$100.00</td>
</tr>
<tr>
<td>130nm</td>
<td>$10.00</td>
</tr>
<tr>
<td>65nm</td>
<td>$1.00</td>
</tr>
<tr>
<td>32nm</td>
<td>$0.10</td>
</tr>
<tr>
<td>16nm</td>
<td>$0.01</td>
</tr>
<tr>
<td>8nm</td>
<td>$0.01</td>
</tr>
<tr>
<td>4nm</td>
<td>$0.01</td>
</tr>
<tr>
<td>2nm</td>
<td>$0.01</td>
</tr>
</tbody>
</table>

Source: Objective Analysis
What Dictates Memory Cost?

- Cost per megabyte depends on:
  - Wafer cost
  - Megabytes per wafer
  - Yield

- Megabytes per wafer driven by bit size
  - Shrinking bits allow cost reductions
  - Manufacturers shrink processes to drive this

This is Moore’s Law in Action!
NAND $/GB > DRAM’s Until 2004

Average Price per Gigabyte

Source: Objective Analysis
In 2004 NAND GB $\frac{1}{3}$ of DRAM’s
The Same is True of All Memory Technologies

There can be no price advantage without comparable scale
New Memory Prices Will Move Past Established Technologies

- New memory migrates from lagging to leading-edge processes
- Technologies with shrinking markets do not migrate to advanced processes
- MRAM/SRAM crossover today
- MRAM/NOR crossover ~2019
- New Memory crosses MRAM/NOR crossover ~2019
- DRAM stops scaling later
- New Memory crosses DRAM ~2 generations later

Process leaders DRAM & NAND cost-reduce to Moore’s Law

Source: Objective Analysis, 2018
Outlook

- Nothing works in a vacuum
  - PM is a part of the greater memory ecosystem
  - The memory market swings wildly
- Foundry processes will have a huge impact
Commodity Price Cycle

Shortage
Prices Stabilize

Manufacturers
Under Invest

Manufacturers
Over Invest

Oversupply
Prices Collapse
Status of Today’s Memory Cycle

- Predicted collapse is well underway
  - Supply-driven overcapacity
  - 3D NAND selling close to cost
  - DRAM still has room to fall
- Won’t end until demand catches up with supply
  - Typically this takes 2 years
  - China likely to enter market in 2021
Impact to PM?

- Persistent memory competes against established technologies
  - Example: 3D XPoint must be cheaper than DRAM
- A DRAM collapse will create an XPoint collapse
  - Even though XPoint is sole-sourced!
Timeline for Change

Logic

NAND

DRAM

Source: Objective Analysis, 2018
Growth in New Memory Shipments

- Embedded MRAM replaces most SoC NOR and SRAM
- Strong appeal in AI apps

Could exceed $37B by 2029!
New Memory Capital Spending Increases

>$800 Million!
New Emerging Memory Report

- Coughlin Associates/Objective Analysis
- Examines Emerging Memory Ecosystem
  - Technologies (PCM, ReRAM, MRAM, FRAM…)
  - Companies
  - Markets
  - Support requirements
- Forecasts Emerging Memory consumption
  - Embedded Emerging Memories
  - Discrete Emerging Memories
- 172 pages, 30 tables, 125 figures

http://www.tomcoughlin.com/techpapers.htm
https://Objective-Analysis.com/reports/#Emerging
Summary

- Emerging Memories Making Good Progress
  - Major commitments for server and embedded applications
  - Still hard to determine the frontrunner
- Support requirements being well addressed
- New memories will drive capital spending
- Many issues confront the market
Emerging Memories Poised to Explode: Emerging Memory Report
Digital Storage in Media and Entertainment
Digital Storage Technology Newsletter
Emerging Memory and Artificial Intelligence Workshop, Stanford, 8/29/19
Storage Valley Supper Club
OBJECTIVE ANALYSIS

Profound Analysts

Reports & Services

Custom Consulting

2019 Storage Developer Conference. © Objective Analysis & Coughlin Associates. All Rights Reserved.
### Objective Analysis

#### Semiconductor Forecast Accuracy

<table>
<thead>
<tr>
<th>Year</th>
<th>Forecast</th>
<th>Actual</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>Zero growth at best.</td>
<td>-3%</td>
</tr>
<tr>
<td>2009</td>
<td>Growth in the mid teens</td>
<td>-9%</td>
</tr>
<tr>
<td>2010</td>
<td>Should approach 30%</td>
<td>32%</td>
</tr>
<tr>
<td>2011</td>
<td>Muted revenue growth: 5%</td>
<td>0%</td>
</tr>
<tr>
<td>2012</td>
<td>Revenues drop as much as -5%</td>
<td>-2.7%</td>
</tr>
<tr>
<td>2013</td>
<td>Revenues increase nearly 10%</td>
<td>4.9%</td>
</tr>
<tr>
<td>2014</td>
<td>Revenues up 20%+</td>
<td>9.9%</td>
</tr>
<tr>
<td>2015</td>
<td>Revenues up ~10%</td>
<td>-0.2%</td>
</tr>
<tr>
<td>2016</td>
<td>Revenues up ~10%</td>
<td>1.1%</td>
</tr>
<tr>
<td>2017</td>
<td>Revenues up ~20%</td>
<td>22%</td>
</tr>
<tr>
<td>2018</td>
<td>Strong start supports 10+% growth</td>
<td>14%</td>
</tr>
<tr>
<td>2019</td>
<td>Semiconductors down -5%</td>
<td>TBD</td>
</tr>
</tbody>
</table>