Panel Discussion

Goal: Discuss upcoming MRAM Applications in Two Axes: Market and Performance

Panelists:
- Mark Webb, President MKW Ventures
- Daniel Worlege, Distinguished RSM, IBM Research
- Terry Torng, Co-Founder, GyrFalcon
- Tetsuo Endoh, Prof. Tohoku Univ
- Tom Andre, VP Engineering, Everspin
- Jean-Pierre Nozieres, CEO, Antaios

Two Domains:
- Market segments – NVM / Unified / SRAM / AI / Auto, etc
- Performance axis – Retention / Endurance / Cost / Scalability

What is the main Market for MRAM in 2024?
What are the Pros/Cons in Performance?
## Panelist Views

<table>
<thead>
<tr>
<th>Target Market</th>
<th>Mark</th>
<th>Daniel</th>
<th>Terry</th>
<th>Endoh</th>
<th>Tom</th>
<th>Jean-Pierre</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Embedded, SOC, NOR/SRAM replacement</td>
<td>• eNVM @ 105C</td>
<td>• Edge AI and IoT, AI for Robot and AI for automotive</td>
<td>• e-Memory (SRAM &amp; e-Flash replacement)</td>
<td>• Entry: Storage (SSD)</td>
<td>• Immediate: IoT (battery-operated, ex. Data logging)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• eNVM @ 125C, solder reflow</td>
<td>• eNVM @ 125C with solder reflow is doable now</td>
<td>• Non volatile, no leakage</td>
<td>• High bandwidth non-volatile writes</td>
<td>• Fast and non volatile</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Mobile cache (10 ns)</td>
<td>• Mobile cache at 10 ns is doable now</td>
<td>• Memory density</td>
<td>• High Endurance</td>
<td>• Huge power savings at chip level</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Last level cache</td>
<td></td>
<td>• Scalability</td>
<td>• Reliability</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Potential universal memory</td>
<td>• ST-DDR4 Interface</td>
<td></td>
</tr>
<tr>
<td>Pros</td>
<td>Ability to support DRAM, NOR, SRAM applications</td>
<td>• Last level cache is very hard – reliable 2 ns write with low current</td>
<td>• SRAM like ns Read/write speed for some application</td>
<td>• &lt;SRAM replacement&gt;</td>
<td>• Density &lt; DRAM</td>
<td>• Endurance-speed-retention trilemma</td>
</tr>
<tr>
<td></td>
<td>storage in metal layers</td>
<td></td>
<td></td>
<td>• Endurance</td>
<td>• Ind/Auto grades not yet avail</td>
<td>• Heavy ECC required</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• eNVM @ 125C with solder reflow is doable now</td>
<td></td>
<td>• Write Power</td>
<td>• Clock Freq &lt; DRAM</td>
<td>• Limited read speed</td>
</tr>
<tr>
<td>Cons</td>
<td>Maturity in embedded applications</td>
<td>• SRAM like ns Read/write speed for some application</td>
<td>• &lt;SRAM replacement&gt;</td>
<td>• Cost</td>
<td>• STT is not RAM compatible as is</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cost/Scaling</td>
<td></td>
<td>• Endurance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Last level cache is very hard – reliable 2 ns write with low current</td>
<td></td>
<td>• Write Power</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• MRAM will never be SCM – no multi-bits</td>
<td></td>
<td>• Cost</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• &lt;e-Flash Replacement&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Retention@HT</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Cost</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Action Needed</td>
<td>Demonstrated volume in production</td>
<td>• SRAM like performance</td>
<td>• Adv.Process Tech, High throughput</td>
<td>• MTJ Scaling and Design to extend temp and improve power/perf/area</td>
<td>• Achieve speed AND endurance at the same time – SOT!</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Achieve Cell size &lt;20 F²</td>
<td>• MLC MRAM</td>
<td>• ECC Tech</td>
<td></td>
<td>• Improve read margin to raise read speed</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Next-gen MRAM, SOT, VCMA</td>
<td>• Testing Tech</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Satoru Araki: Final Panel 2019 on MRAM in 2024
2019 MRAM Markets and Applications

Mark Webb

MKW Ventures Consulting, LLC
8/5/2019
MRAM Future and Challenges

- MRAM technology is here today
  - We know what it is and what the challenges are
  - Cost, Performance, Density, Endurance, SOC integration, etc
- In past year, we have updates on multiple fronts
  - Existing companies announced 1Gbit parts, updated us on revenue and growth
  - Multiple Companies presented embedded MRAM technologies
    - IEDM/ISSCC Papers were popular. Embedded is an option to choose
  - New technologies and models and optimization
- No need to speculate on what is coming

Mark Webb, MKW Ventures Consulting LLC, 08/05/19
A Tale of Two Markets

- **Embedded**: MRAM is ideal for market (revenue not measurable)
  - Potential to replace NOR, SRAM, DRAM applications
  - Ability to integrate (metal stacks), density (Mbit), performance (DRAM) match embedded needs well
  - Endurance work needed for full RAM replacement
  - Looks like Embedded “Universal Memory”… Why isn’t it here already?

- **Discrete**: Targets and Market growth uncertain (Rev <$100M today)
  - Target applications requires specific density, speed, with NVM requirement.
  - Small markets exist, but they are vulnerable to attack on all sides.
  - No measurable NAND replacement market (Too small, expensive)
  - Performance/cost/density ratio not on track to match DRAM
  - Needs to dominate Niche or have “Killer App”
Revenue Projections for MRAM

- In 2018 we predicted >$900M in Revenue by 2024
  - This will not happen
  - 2018/2019 did not breakout like we hoped/expected

<table>
<thead>
<tr>
<th>Year</th>
<th>MRAM Revenue Baseline</th>
<th>Notes/required milestone</th>
</tr>
</thead>
<tbody>
<tr>
<td>2020</td>
<td>$115M</td>
<td>1Gb selling for revenue in 2020, DRAM-Like performance. Multiple IP sources for foundries</td>
</tr>
<tr>
<td>2022</td>
<td>$217M</td>
<td>Multiple foundries and 1+ Memory company in volume</td>
</tr>
<tr>
<td>2024</td>
<td>$429M</td>
<td>2+ memory companies in volume</td>
</tr>
</tbody>
</table>

Included discrete chips (may be stacked) and revenue from licensing
Does not include embedded memory (no revenue model)

Mark Webb, MKW Ventures Consulting LLC, 08/05/19
What is Needed to meet Revenue by 2024 (Forecast=$429M)

- **ALL MRAM**: Volume production in applications in 2020
  - These are required to allow people to commit the technology to **significant** products.

- **Embedded**: Multiple foundry support with multiple applications
  - MRAM penetration into market is measureable in 2020
  - MRAM becoming chosen technology in 2022 designs

- **Discrete**: Meet aggressive endurance goals, Cell size Goals, Density roadmaps and Cost
  - To have any penetration into larger markets, confidence needs to increase in delivering specs.
  - MRAM is unique enough to not allow easy “backup plans”, so high confidence is needed
More details on MRAM vs other memories in my memory update on Tuesday

Mark Webb, www.mkwventures.com
MRAM Developer Day

Panel Discussion

Daniel Worledge
IBM Research
Senior Manager, MRAM
# Key Advances in Spin-Transfer-Torque MRAM

<table>
<thead>
<tr>
<th>Device</th>
<th>Write</th>
<th>Read</th>
<th>Scaling</th>
</tr>
</thead>
<tbody>
<tr>
<td>1974 Slonczewski (IBM) invents magnetic tunnel junction.</td>
<td>1995 Moodera (MIT) and Miyazaki (Tohoku U.) demonstrate first room temperature magnetic tunnel junctions.</td>
<td>1996 Slonczewski (IBM) invents spin-transfer-torque switching.</td>
<td>2004 Parkin (IBM) and Yuasa (AIST) publish discovery of high magnetoresistance in MgO tunnel junctions.</td>
</tr>
<tr>
<td>2010 Worledge (IBM) and Ohno (Tohoku U.) demonstrate first perpendicular CoFeB tunnel junctions.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Magnetic Tunnel Junction**
- **Low Resistance**
- **High Resistance**

**Spin Transfer Torque**
- |<| I_c|
- |≥| I_c|
- |≈| I_c|

**MgO Tunnel Barriers**

**Perpendicular Magnetization**
MRAM Applications

### Standalone
- Replace battery-backed SRAM or DRAM
- Buffer for hard disk drive
- Replace DRAM
- Lower temp process is OK
- 0 C – 70 C operation
- 256 Mb – 1 Gb and up
- 30 - 70 ns read/write
- High endurance ($10^{10} - 10^{15}$)

### Embedded Non-volatile
- Replace NOR eFlash to store:
  - Microcontroller code
  - Encryption key storage
  -Trimming and calibration
- 400C process required
- -40 C – 105/125/150 C
- 1 – 64 Mb
- 30 ns read, 200 ns write
- Low endurance: $10^6$

### Mobile Cache
- Replace SRAM for low performance & power apps
- Wearable electronics
- Co-processors
- Internet of Things
- 400C process required
- 0 C – 85 C operation
- 1 Gb and up
- 1 - 2 ns read/write
- Unlimited endurance ($10^{18}$)

### Last Level Cache
- Fast dense memory for L3 or L4 cache
- Alternative to eDRAM
- 400C process required
- 0 C – 85 C operation
- 1 Gb and up
- 1 - 2 ns read/write
- Unlimited endurance ($10^{18}$)
MRAM in Edge AI and AIoT

Terry Torng

*Stealth Startup
co-founder
Gyrfalcon Technology Inc
Core Challenge To AI: Energy Efficiency

- **Data Center Energy Use is Growing….**
  - “Global data centers used roughly 3% of total electricity in 2016, and will **double** every four years”
  - Radoslav Danilak, December 15, 2017

- **Edge and IoT Devices….**
  - “AI is hungry for processing power. IoT is projected to exceed **20b devices by 2020**. There are currently 10b internet-connected devices, doubling to 20 billion will require massive increases to our data center infrastructure, which will massively increase our **electricity consumption**.” Radoslav Danilak, December 15, 2017

- “Global IP traffic will increase nearly threefold over the next five years, and will have increased 127-fold from 2005 to 2021.”
  - Bill Kleyman, Mar 09, 2018

---

Santa Clara, CA
August 2019
MRAM for AI

Challenges for the Edge AI

- Power consumption
- Application specific
- Algorithm and model
- Data management and data flow control

Memory: 75 to 80% of chip area
Memory: 67% or more leakage power

SRAM like MRAM:
1. Few ns read/write,
2. $10^{16}$ endurance,
3. Memory density 3:1 or more vs SRAM.

Goal:
1. Chip size ½ of less.
2. AIoT (low duty), 60% less power consumption than SRAM only solution.
3. Less than 1mw “high performance” AI accelerator for edge or AIoT
PPP (power, performance and price) Now, Next and Future

- Hardware/software/MRAM co-design
- Simplify circuit design
- Manufacturing friendly for foundries
  - high DR/R materials and different thermal budget processes…
- OST, SOT, voltage-controlled and MLC MRAM compatible
- Chip/wafer yield friendly
What is the main segment for MRAM in 2024?

> What is the main segment for MRAM in 2024?

What is the main segment for MRAM in 2024?

<table>
<thead>
<tr>
<th>CMOS SRAM-LLC</th>
<th>MTJ STT-MRAM(S)</th>
<th>CMOS eFlash</th>
<th>MTJ STT-MRAM(F)</th>
<th>Other NVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operation Voltage</td>
<td>&lt;1.1 V</td>
<td>&lt;0.5 V</td>
<td>12 V</td>
<td>&lt;0.5V</td>
</tr>
<tr>
<td>Write Current</td>
<td>10^-5 A</td>
<td>10^-5 A</td>
<td>10^-4 A</td>
<td>10^-5 A</td>
</tr>
<tr>
<td>Write Speed</td>
<td>&lt;10ns</td>
<td>&lt;10 ns</td>
<td>10000 ns</td>
<td>&lt;200 ns</td>
</tr>
<tr>
<td>Read Speed</td>
<td>&lt;5ns</td>
<td>&lt;5 ns</td>
<td>30 ns</td>
<td>&lt;25 ns</td>
</tr>
<tr>
<td>Retention</td>
<td>Volatile</td>
<td>1～Several Month</td>
<td>10Years</td>
<td>10Years</td>
</tr>
<tr>
<td>Endurance</td>
<td>10^15</td>
<td>10^15</td>
<td>10^3~4</td>
<td>&lt; 10^8~12</td>
</tr>
<tr>
<td>Cell Size</td>
<td>160 ~ 280F^2</td>
<td>12 ~ 28F^2</td>
<td>64 ~ 128F^2</td>
<td>6 ~ 14F^2</td>
</tr>
</tbody>
</table>

- Low Power
- Small Cell Size

Which application / market will drive it?

> Which application / market will drive it?

- Mobile, IoT, Car Electronics, AI, Robot etc

<table>
<thead>
<tr>
<th>CMOS SRAM-LLC</th>
<th>MTJ STT-MRAM(S)</th>
<th>CMOS eFlash</th>
<th>MTJ STT-MRAM(F)</th>
<th>Other NVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operation Voltage</td>
<td>&lt;1.1 V</td>
<td>&lt;0.5 V</td>
<td>12 V</td>
<td>&lt;0.5V</td>
</tr>
<tr>
<td>Write Current</td>
<td>10^-5 A</td>
<td>10^-5 A</td>
<td>10^-4 A</td>
<td>10^-5 A</td>
</tr>
<tr>
<td>Write Speed</td>
<td>&lt;10ns</td>
<td>&lt;10 ns</td>
<td>10000 ns</td>
<td>&lt;200 ns</td>
</tr>
<tr>
<td>Read Speed</td>
<td>&lt;5ns</td>
<td>&lt;5 ns</td>
<td>30 ns</td>
<td>&lt;25 ns</td>
</tr>
<tr>
<td>Retention</td>
<td>Volatile</td>
<td>1～Several Month</td>
<td>10Years</td>
<td>10Years</td>
</tr>
<tr>
<td>Endurance</td>
<td>10^15</td>
<td>10^15</td>
<td>10^3~4</td>
<td>&lt; 10^8~12</td>
</tr>
<tr>
<td>Cell Size</td>
<td>160 ~ 280F^2</td>
<td>12 ~ 28F^2</td>
<td>64 ~ 128F^2</td>
<td>6 ~ 14F^2</td>
</tr>
</tbody>
</table>

- Low Power
- High Speed Write
- Better Endurance
- Small Cell Size
What are the challenges?

<My Position>

**Improvement of MRAM Performance for each Application**
- For higher level cache: High Speed Write & Low Write Power & over $10^{15}$ Endurance@128Mb-GbMRAM
- For e-Flash of Automotive Application: Excellent Thermal Stability ($\Delta$)
- For Main Memory: Scalability & High Density Array

What technological breakthrough needed?

<My Position>

1) Advanced Process Technology including High throughput & Damage Control Technology
2) Circuit Technology including ECC Technology
3) Testing Technology for mass production

Most critical problem of MRAM is our mind.
⇒ **Tough Mind** to believe MRAM Potential
MRAM Developer Day 2019
Panel: MRAM Application – Pros/Cons

Tom Andre
VP Engineering, Everspin Technologies
STT-MRAM Applications Today

- Larger Buffer Improves QOS
- Enable Higher Number of Streams
- Simplified Architecture Eliminates Power Fail Hardening
- More Physical Space For Storage Capacity
- No Capacitor Liability
- Optimized Interleave For Sequential Performance

![Graph showing QoS improvement](image)

- 25% - 45% latency improvement
- Latency improvement is architecture dependent
## MRAM Performance

<table>
<thead>
<tr>
<th>Features</th>
<th>16Mb Toggle</th>
<th>256Mb ST-DDR3</th>
<th>1Gb ST-DDR4</th>
</tr>
</thead>
<tbody>
<tr>
<td>VDD / VPP</td>
<td>3.3V</td>
<td>1.5V</td>
<td>1.2V / 2.5V</td>
</tr>
</tbody>
</table>
| Data Retention | 20 years / 85°C  
2 years / 125°C | 3 months / 70°C  
3 months / 70°C | 10 yrs / 85°C capable |
| Endurance | $>1e16$  
($>10$ yrs constant use) | $1e10$ cycles  
every page | $1e10$ cycles  
every page |
| Uniform 10 yr writes per Chip | 18 PBW  
(performance limit) | 320 PBW  
(cycle limit) | 840 PBW  
(performance limit) |
| Peak Bandwidth per x16 Chip | 57MB/s | 2.67GB/s | 2.67GB/s |
## STT-MRAM Areas for Development

<table>
<thead>
<tr>
<th></th>
<th>Storage / SSD</th>
<th>Industrial</th>
<th>Auto</th>
<th>Aerospace</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Application</strong></td>
<td>Non-volatile Data Buffer/Cache/PMR</td>
<td>Non-volatile Data logs, params</td>
<td>Non-volatile Data params, buffer</td>
<td>Non-volatile Code, Data</td>
</tr>
<tr>
<td><strong>STT-MRAM benefit</strong></td>
<td>Write Bandwidth, Endurance</td>
<td>Endurance, Reliability</td>
<td>Write BW, End, Reliability</td>
<td>Rad Hard, Reliability</td>
</tr>
<tr>
<td><strong>Density</strong></td>
<td>128MB - &gt;1GB</td>
<td>8MB - 128MB</td>
<td>128MB - &gt;1GB</td>
<td>8MB - 128MB</td>
</tr>
<tr>
<td><strong>Operating Temp</strong></td>
<td>0C - 85C</td>
<td>-40C – 85C</td>
<td>-40C – 125C (Grade 1)</td>
<td>-40C – 125C</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-40C – 150C (Grade 0)</td>
<td>-55C – 125C</td>
</tr>
<tr>
<td><strong>Data Retention</strong></td>
<td>3 months / 70C</td>
<td>10 years / 85C</td>
<td>Profile across temp</td>
<td>15 yrs / 105C</td>
</tr>
<tr>
<td><strong>Endurance</strong></td>
<td>Buffer: 1e10 every page PMR: &gt;1e15 total</td>
<td>varies by app &gt;1e15 total</td>
<td>varies by app</td>
<td>varies by app</td>
</tr>
<tr>
<td><strong>Peak Bandwidth</strong></td>
<td>5GB/s - &gt;12GB/s</td>
<td>10MB/s – 2GB/s</td>
<td>10MB/s – 2GB/s</td>
<td>&lt;100MB/s</td>
</tr>
</tbody>
</table>
Mission Statement: Jointly develop and license SOT (technology and IPs) as the next-generation MRAM that solves STT shortcomings.

Endurance-Retention-Speed tradeoff

- Limited to e-NVM applications

Infinite endurance / High intrinsic speed

= Fully RAM compatible
(ensures Cache applications)
PERFORMANCES PROS & CONS (STT)

Pros

- Fast and low (write) power, beats eFlash hands down
  - Huge power savings at chip level (+ endurance)
- Fast and non volatile
  - Zero standby power, zero leakage, unlike SRAM

Cons (STT)

- STT not RAM compatible
  - Endurance / speed / retention trilemma
  - Low read window = limited read speed
  - Heavy ECC needed (cost, performance limitations)
MRAM APPLICATIONS (STT)

“Low Hanging Fruit” Market: IoT

- In particular for those applications involving large amounts of data for which power and endurance are key (ex. Data logging, audio/video monitoring)

Other “near-ready” markets

- **Automotive** - Provided the temperature/reliability specs are met
- **AI chips** - Killer benefits of MRAM (can hold data during compute)
  
  STT good enough as-is, but SOT should be even better

STT is excellent for NVM (e-Flash) replacement, but full benefit of MRAM migration will be achieved when cache (e-SRAM) is addressed as well
To unleash full potential of embedded MRAM ...

- **Improve device towards RAM requirements**
  - Need speed AND endurance at the same time
    - STT won’t do it - Need something new
    - SOT is the only solution (fast write & infinite endurance)
  - Improve read margin to raise (read) speed
    - By materials (TMR increase)
    - By design (device and chip level)

To address standalone market ...

- **Achieve high capacities (>Gb)**
  - Control of BER, distribution (tails), ...
  - Dedicated memory process / architecture for high density (low F2)