Nanotechnology Trends in Nonvolatile Memory Devices

Gian-Luca Bona

gianni@us.ibm.com
IBM Research,
Almaden Research Center
The Elusive Universal Memory

"Universal" memory market to hit $75 billion in 2019, says iSuppli

Peter Clarke
(07/27/2005 7:45 AM EDT)
URL: http://www.eetimes.com/showArticle.jhtml?articleID=166402857

LONDON — The market for a memory integrated circuit that combines the speed of SRAM, the density of DRAM and the non-volatility of flash, could be $76.3 billion by 2019, according to market research company iSuppli Corp. (El Segundo, Calif.).

The so-called “universal” memory would, by then, have grabbed about 80 percent of the market, the market researcher has estimated in a long-range forecast it described as “speculative.”

There is no single semiconductor memory technology today that has all the desired attributes, which on top of speed, density and non-volatility include: low-cost of manufacture, low switching energy and scalability to nanometer-scale dimension.

Products in various stages of commercialization that include at least some of the attributes include: Ovonic Unified Memory (OUM), Magneto-Resistive RAM (MRAM), Ferroelectric RAM (FRAM) and Nanotube RAM (NRAM), iSuppli said. But the rewards for a winning technology are likely to be immense with the memory market set to double from $46.8 billion posted in 2004 to $95.4 billion by 2019, iSuppli said.

The market researcher said that it does not usually forecast markets beyond a five-year horizon. However, due to the emerging status of the universal memory market, a longer-range forecast is required.
Incumbent Semiconductor Memories

Attributes for universal memories:
- Highest performance
- Lowest active and standby power
- Unlimited Read/Write endurance
- Non-Volatility
- Compatible to existing technologies
- Continuously scalable
- Lowest cost per bit
Incumbent Semiconductor Memories

A new class of universal storage device:

- a fast solid-state, nonvolatile RAM
- enables compact, robust storage systems with solid state reliability and significantly improved cost-performance
Non-volatile, universal semiconductor memory

- Everyone is looking for a dense (cheap) crosspoint memory.
- It is relatively easy to identify materials that show bistable hysteretic behavior (easily distinguishable, stable on/off states).
The Memory Landscape
Histogram of Memory Papers

Emerging Memory Technologies
Memory technology remains an active focus area for the industry

<table>
<thead>
<tr>
<th>FLASH Extension</th>
<th>FRAM</th>
<th>MRAM</th>
<th>PCRAM</th>
<th>RRAM</th>
<th>PCM-SS Electrolyte</th>
<th>Polymer/Organic</th>
<th>Mechanical</th>
<th>3D</th>
<th>Thyristor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trap Storage</td>
<td>Ramtron</td>
<td>IBM</td>
<td>Ovonix</td>
<td>IBM</td>
<td>Axon</td>
<td>Spansion</td>
<td>Nantero</td>
<td>Matrix</td>
<td>T-RAM</td>
</tr>
<tr>
<td>Saifun NROM</td>
<td>Fujitsu</td>
<td>Philips</td>
<td>BAE</td>
<td>Sharp</td>
<td>Samsung</td>
<td>STMicro</td>
<td>Hitachi</td>
<td>Sony</td>
<td></td>
</tr>
<tr>
<td>Tower</td>
<td>STMicro</td>
<td>Infineon</td>
<td>Freescale</td>
<td>Unity</td>
<td>Infineon</td>
<td>Quimonda</td>
<td>MEC</td>
<td>3D-ROM</td>
<td></td>
</tr>
<tr>
<td>Spansion</td>
<td>TI</td>
<td>STMicro</td>
<td>STMicro</td>
<td>Fliolda</td>
<td>Infineon</td>
<td>Hitachi</td>
<td>Zettagore</td>
<td>Samsung</td>
<td></td>
</tr>
<tr>
<td>Infineon</td>
<td>Toshiba</td>
<td>Samsung</td>
<td>Samsung</td>
<td>Infineon</td>
<td>Philips</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Macronix</td>
<td>Infineon</td>
<td>HP</td>
<td>Infineon</td>
<td>Samsung</td>
<td>Philips</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Samsung</td>
<td>Samsung</td>
<td>NVE</td>
<td>Infineon</td>
<td>Samsung</td>
<td>Philips</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Toshiba</td>
<td>NEC</td>
<td>Honeywell</td>
<td>NVE</td>
<td>Infineon</td>
<td>Philips</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanson</td>
<td>Hitachi</td>
<td>Toshiba</td>
<td>Infineon</td>
<td>Philips</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Macronix</td>
<td>Rohm</td>
<td>NEC</td>
<td>Infineon</td>
<td>Philips</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEC</td>
<td>HP</td>
<td>Sony</td>
<td>Infineon</td>
<td>Philips</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nano-x’tal</td>
<td>Cypress</td>
<td>Fujitsu</td>
<td>Sony</td>
<td>Infineon</td>
<td>Philips</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Freescale</td>
<td>Matsushita</td>
<td>Renesas</td>
<td>Samsung</td>
<td>Infineon</td>
<td>Philips</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Matsushita</td>
<td>Oki</td>
<td>Hynix</td>
<td>Samsung</td>
<td>Infineon</td>
<td>Philips</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

IBM working towards a 16GB part by 2010

STMicroelectronics is claiming significant progress in the development of a new type of electronic memory that could eventually replace Flash memory technology.

<table>
<thead>
<tr>
<th>PERFORMANCE</th>
<th>DURABILITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Flash</td>
<td>Nonvolatile</td>
</tr>
<tr>
<td>High SSD</td>
<td>Volatile</td>
</tr>
<tr>
<td>N/A</td>
<td>DRAM</td>
</tr>
</tbody>
</table>

4Mb C-RAM (Product)
0.25um 3.3V

2Mb FRAM (Product)
0.35um 3.3V

4Mb MRAM (Product)
0.18um 3.3V

512Mb PRAM (Prototype)
0.1um 1.8V
Critical applications are undergoing a paradigm shift

**Compute-centric paradigm**

- Solve differential equations
- CPU / Memory
- Computational Fluid Dynamics
- Finite Element Analysis
- Multi-body Simulations

**Data-centric paradigm**

- Analyze petabytes of data
- Storage & I/O
- Search and Mining
- Analyses of social/terrorist networks
- Sensor network processing
- Digital media creation/transmission
- Environmental & economic modeling

**Main Focus:** Bottleneck

**Typical Examples:**

- Disks or Flash can’t keep up with data-centric applications
- Thesis: Disks or Flash can’t keep up with data-centric applications
- Proposal: Develop device technology and build a high-density array and demonstrate performance and endurance for the data-centric paradigm
What are the limitations with disks?

- Bandwidth – Access Time – Reliability - Power

- Disk Performance improves very slowly
  - Gap between processor and disk performance widens rapidly

- Bandwidth 100MB/s – slow improvement
  - gap can be solved with many parallel disks
  - but need 10,000 disks today, >1,000,000 disks by 2020
    - but that’s just for a traditional high-end HPC system
    - data intensive problems are much worse

- Access time gap has no good solution
  - disk access times (msec); decrease only 5% per year
  - complex caching or task switching schemes help - sometimes

- Disk power dissipation is a major factor in data-centric systems (~4W/disk)
- Newest disk generations are less reliable than older ones
  - Data losses occur in even the best enterprise-class storage systems
Power & space in the server room

The cache/memory/storage hierarchy is rapidly becoming the **bottleneck for large systems**.

We know how to create MIPS & MFLOPS cheaply and in abundance, but feeding them with data has become the performance-limiting and most-expensive part of a system (in both $ and Watts).

---

**U.S. Market**

- New server spending
- Power and cooling

**Extrapolation to 2020**

(at 90% CGR → need 1.7 PB/sec)

- 5.6 million HDD
  - 19,000 sq. ft. !!
  - 25 Mega watts

---


---

What are the limitations with Flash?

- **Read/Write Access Times** – Write endurance – Block architecture

- **Flash Performance showing no improvement**
  - Gap between processor and Flash performance continues to widen
  - Write endurance \(<10^6\) and showing no improvement trends
    - Need \(>10^9\) to cater to frequent writes as data continually flows into the system
      - Tomorrow’s hand-held devices will be continuously updated
      - Intel applications characterized by continuous data streams
  - Access time gap has no good solution
Processing Cost and $F^2$

- The bit cell size drives the cost of any memory
- Cell area is expressed in units of $F^2$ where $F$ is the minimum lithographic feature of the densest process layer
  - Half pitch dimension of metallization connecting drain and source for ICs
  - MR sensor width in magnetic recording
- Cell areas
  - DRAM: $8F^2 \rightarrow 6F^2$
  - NAND: $4F^2 \rightarrow 2F^2$
  - SRAM: $100F^2$
  - MRAM: $15F^2 -- 40F^2$
  - Hard Disk: $0.5F^2 \rightarrow 1F^2$
Density

What’s next after CD reaching Physical Limit?

- Beyond the lithographic CD limit, there are 2 ways to continue Moore’s Law of Cost Reduction in Semiconductor Memories:
  - Multi-bit per cell (MLC),
  - Multi-layer stacking (3D).

- Multi-bits per cell is the more effective way, the combination is most powerful:
  - 8-layer stack is probably the cost-effective limit for fully integrated stacking,
  - 2 bits per cell is probable with Phase Change Memories,
  - >2 bits would require more innovation.
Storage Historic Price Trend and Forecast

- 34nm 4-Layer 1-bit
- 27nm 4-Layer 2-bit
- 22nm 4-Layer 2-bit
- 18nm 4-Layer 3-bit
- 14nm 4-Layer 3-bit
- 14nm 4-Layer 4-bit

Ref. C. Lam

© 2008 IBM Corporation
Universal Memory or Storage Class Memory Target Specifications

<table>
<thead>
<tr>
<th>Feature</th>
<th>Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access Time</td>
<td>~100-200 ns</td>
</tr>
<tr>
<td>Data Rate (MB/s)</td>
<td>100</td>
</tr>
<tr>
<td>Endurance</td>
<td>$10^9 - 10^{12}$</td>
</tr>
<tr>
<td>HER (/TB)</td>
<td>$10^{-4}$</td>
</tr>
<tr>
<td>MTBF (MH)</td>
<td>2</td>
</tr>
<tr>
<td>On Power (mW)</td>
<td>100</td>
</tr>
<tr>
<td>Standby (mW)</td>
<td>1</td>
</tr>
<tr>
<td>Cost ($/GB)</td>
<td>&lt;5.5</td>
</tr>
<tr>
<td>CGR</td>
<td>35%</td>
</tr>
</tbody>
</table>

Very challenging to achieve in combination
SCM Basic Concepts: Phase Change Example

- Using a phase transition of a *Ge-Sb-Te alloy to store a bit*
- *Ge-Sb-Te* exists in a stable amorphous and a stable crystalline phase
  - Phases have very different electrical resistances
- **Transition between phases by controlled heating/cooling**
  - Write ‘1’: short (10ns) intense current pulse melts alloy crystal => amorphous
  - Write ‘0’: longer (50ns) weaker current pulse re-crystalizes alloy => crystalline
  - Read : short weak pulse senses resistance, but doesn’t change phase
- **Issue**: rectifying diode materials for high-ON current density (> $10^7$ A/cm² – needed for PCM) and ultra-low OFF current density (< 1 A/cm²).
A Brief History of Phase Change Memory

1968 Ovshinsky published Phase Change Threshold Switching
1969
1970 R.G. Neekle et al demonstrated a 250-nm Phase Change Memory
1971
1972 J. Feinleib et al demonstrated Reversible Optical Memory
1974
1978
1981
1982
1986
1987
1988
1989
1990
1991
1994
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007

IBM Research

© 2008 IBM Corporation

We describe here a rapid and reversible transition between a highly resistive and a conductive state effected by an electric field which we have observed in various types of disordered materials, particularly amorphous semiconductors covering a wide range of compositions. These include oxide- and boron-based glasses and materials which contain the elements tellurium and/or arsenic combined with other elements such as those of groups III, IV, and VI.

Such amorphous materials can be described as intrinsic semiconductors with an optical energy gap $E_g$ typically between 0.6 and 1.4 eV and an activation energy $E_a$ for electrical conduction $E_a$ between 0.7 and 1.9 eV depending on composition.

The experimental setup indicates that the doping (in atomic percent) 48 at. % tellurium, 30 at. % arsenic, 12 at. % silicon, and 10 at. % germanium. The specimen was evaporated film, $5 \times 10^{-3}$ cm thick, between two carbon electrodes with a contact area of about $10^{-4}$ cm$^2$. This material has a resistivity at $300^\circ$K of $3 \times 10^3 \Omega$ cm, $AE$ $1.0$ eV, and a positive temperature.

Figure 1 shows oscilloscope pictures of (a) the $I$-$V$ characteristic, (b) the voltage $V$ across the unit, and (c) the current $I$ passing through the above unit as a function of time. In this case, a 60-Hz ac voltage was applied across the unit and a 10-ohm load resistor was used. The $I$-$V$ curve is independent of frequency to at least 10$^5$ Hz.

The major features of the switching phenomena shown are the following: (1) The $I$-$V$ characterist-

![Image]

**FIG. 1.** Response of switching unit to 60-Hz voltage. (a) $I$-$V$ characteristic: vertical, 2 mA/div; horizontal, 5 V/div. (b) Voltage: vertical, 5 V/div horizontal, 5 msec/div. (c) Current: vertical, 20 mA/div; horizontal, 5 msec/div.
A Long “Pause”

1962 A.D. Pearson et al reported switching phenomena in ArTel (Advanc. Galss Tech. p357)
1968 Ovshinsky published Phase Change Threshold Switching
1969
1970 R.G. Neale et al demonstrated a 256-bit Phase Change Memory
1971
1972 J. Feinleib et al demonstrated reversible Optical Memory
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990 Panasonic introduced R/W Phase Change Optical Disk Drive
1991
1992
1993
1994
1995
1996
1997
1998
1999 Ovonyx formed
2000 BAE licensed Ovonic Unified Memory (OUM) from Ovonyx 11/4/1999
2000 Intel Capital invested in Ovonyx and licensed OUM from Ovonyx 2/9/2000
2001 STMicro licensed OUM from Ovonyx and announce Joint Development Project 12/21/2000
2002
2003 STMicro and Ovonyx expanded scope of OUM license and extended IUP 7/4/2003
Samsung published first paper on PRAM in HLST’03
Mitsui published first paper on PRAM in IBM’03
2004 Samsung announced full-scale PRAM production in 2005/02/20/2004
Samsung announced full-scale PRAM production in 2005/02/20/2004
Nanochip licensed OUM from Ovonyx for Micro-Electro_Mechanical Systems (MEMS) based storage 8/31/2004
2005 Elpida licensed OUM from Ovonyx 2/3/2005
2005 Phillips research published first paper on PRAM in Nature Materials April issue
Samsung announced 512Mb PRAM availability in 2006 to replace VCR Flash 12/28/2005
2006 ITRI of Taiwan and local memory vendors - PowerChip, Nanya, Promos and Winbond formed PCM Alliance 9/28/2006
2007 Ovonyx and Qimonda Sign Technology Licensing Agreement for Phase Change Memory 1/10/2007

256 bits 25V 7.5mA 15ms, 25V 150mA 6us

Intel, ECD 1970

The energy required to melt the Phase Change Memory Element scales with CD ...
Phase-Change Nano-Bridge
- Prototype memory device with ultra-thin (3nm) films demonstrated Dec '06

- $3\text{nm} \times 20\text{nm} \rightarrow 60\text{nm}^2$
- $\approx$ Flash roadmap for 2013
  $\rightarrow$ phase-change scales

- Fast (<100ns SET)
- Low current (< 100$\mu$A RESET)

$W$ defined by lithography
$H$ by thin-film deposition

Current scales with area
- $H=25\text{nm}$
- $H=3\text{nm}$
- $H=10\text{nm}$
- $L=50\text{nm}$

$\text{SET current (\mu A)}$
$\text{RESET current (\mu A)}$
$\text{Time (ns)}$

Area = $W \times H$ (in nm$^2$)
Crossbar Memory Fundamentals

standard crossbar memory

Cell size = 4F^2

What if we can put more cells at a crossbar?

Net effect: Density ↑ n^2
Cost ↓ n^2

1-D

4F^2/n

Memory Cells between CMOS lines

2-D

4F^2/n^2
• Sub lithographic feature is selected by moving depletion across the fine structure
• Modulating signal is brought in by lithographically defined lines
• Fins down to sub 20 nm have been addressed
MNAB Concept Demonstrated

100nm Pitch MNAB Devices Fabricated by E-Beam Lithography

Obtained Fully Functional Devices

Selectivity > $10^5$

Gate 1 = -2.0V

Fin Current (A)

Gate 2 Voltage (V)

Fin Currents (A)

Gate 2 Voltage

Unselected Fins (1 + 3 + 4)

Selected Fin (2)

Selected Fin (1)

Unselected Fins (2 + 3 + 4)

Corporation
Combining Micro-Nano Decoder and ROM

4-fin UMB+ROM test structure

- Successful integration of UMB with memory element
  (2 terminal oxide antifuse ROM)
- Verified operation over all bit sequences for 4-fin UMB+ROM
Nanoscale Patterning Techniques

- Litho Tool: 193nm immersion at 1.35 NA, next?
- Various nanoscale patterning techniques exist.
- Simple regular line / space patterns possible.

IBM Research
Self Assembly

IBM Lithography

Spacers
Frequency doubling – 40 nm to 20 nm pitch (IBM)

Nanolithography
Princeton / Nanonex
**Step-and-Flash Imprint Lithography (SFIL)**

---

**Silicon Fins Resulting from Oxide and Imprint Etch Masks**

- Critical Dimension Control
- Side-Wall Profile
- Line-Edge Roughness

---

**Silicon Fins Ready for Ion-Implant Lithography and Processing**

Mark Hart, et al.
Mix-and-Match Overlay of SFIL to Optical Lithography

-- To Align Optical Levels with Imprint Level --

Generate “Zero-Level” Marks in Wafer via 193nm Lithography and Etch to Ensure they Survive the Full MNAB Process Build

Align both Imprint Level and Subsequent Optical Levels to These Zero-Level Marks

-- Using This Approach --

Demonstrated Sub-20nm (Mean+3σ) Overlay Between 193nm Litho Zero-Level and Imprint Over Full 200mm Wafers

Routinely Achieving Sub-50nm Overlay in Approximately 75% of Fields
Paths to ultra-high density memory

starting from standard $4F^2$ ...

...store $M$ bits/cell with $2^M$ multiple levels

...add $N$ 1-D sub-lithographic “fins” ($N^2$ with 2-D)

demonstrated (at IEDM 2005)

...go to 3-D with $L$ layers

demonstrated (at IEDM 2007)
Multi-level phase-change memory

10x10 test array

16k-cell array

Write Strategies for 2 and 4-bit Multi-Level Phase-Change Memory
T. Nirschl\textsuperscript{1,2}, J. B. Philipp\textsuperscript{3}, T. D. Happ\textsuperscript{4}, G. W. Burg\textsuperscript{5}, B. Rajendran\textsuperscript{1}, M.-H. Lee\textsuperscript{2}, A. Schrott\textsuperscript{1}, M. Yang\textsuperscript{2}, M. Brettwisch\textsuperscript{1}, C.-F. Chen\textsuperscript{2}, E. Joseph\textsuperscript{1}, M. Lamo\textsuperscript{1}, R. Cheek\textsuperscript{1}, S.-H. Chen\textsuperscript{1}, S. Zaikin\textsuperscript{2}, S. Raoux\textsuperscript{1}, Y.C. Chen\textsuperscript{1}, Y. Zhu\textsuperscript{1}, R. Bergmann\textsuperscript{1}, H.-L. Lung\textsuperscript{2}, C. Lam\textsuperscript{1}

IBM/Qimonda/Macronix PCRAM Joint Project

IBM T.J. Watson Research Center, 1101 Kitchawan Road, Yorktown Heights, NY, 10598, USA,
\textsuperscript{2}Infineon Technologies, \textsuperscript{1}Qimonda, \textsuperscript{1}IBM Yorktown, \textsuperscript{1}IBM Essex Junction, \textsuperscript{1}IBM Almaden, \textsuperscript{2}Macronix International Co. Ltd.

IEDM 2007
Paths to ultra-high density memory

At the 32nm node in 2013, MLC NAND Flash (already $M=2 \rightarrow 2F^2$ !) is projected* to be at...

<table>
<thead>
<tr>
<th>Density</th>
<th>Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2x$ 43 Gb/cm²</td>
<td>32GB</td>
</tr>
<tr>
<td>$4x$ 86 Gb/cm²</td>
<td>64GB</td>
</tr>
<tr>
<td>$16x$ 344 Gb/cm²</td>
<td>256GB</td>
</tr>
<tr>
<td>$64x$ 1376 Gb/cm²</td>
<td>$\sim$1 TB</td>
</tr>
</tbody>
</table>

if we could shrink $4F^2$ by...

e.g., 4 layers of 3-D ($L=4$)
e.g., 8 layers of 3-D, 2 bits/cell ($L=8, M=2$)
e.g., 4 layers of 3-D, 4x4 sublithographic ($L=4, N=4^2$)

* 2006 ITRS Roadmap
Magnetic Racetrack Memory: a 3-D shift reg. Memory

- Data stored as pattern of domains in long nanowire or “racetrack” of magnetic material.
- Data stored magnetically and is non-volatile.
- Current pulses move domains along racetrack - *no moving parts, just the patterns move.*
- Each memory location stores *an entire bit pattern* (10, 100, 1000 bits?) rather than just a single bit.
Magnetic Racetrack Memory Concept

Current Pulse Train Drives Domains

Racetrack Engineered
So Domains “Stick” at precise intervals

Domains Move Around Track

“Storage” at each end so entire pattern can move past heads in either direction

As Domain Walls Pass MTJ “head”, Data is Read Out

Current Pulse Here Can Write New Data
Magnetic Race-Track Memory

- Information stored as domain walls in vertical “race track”
  - Data stored in the third dimension in tall columns of magnetic material
- Domains moved around track using nanosecond pulses of current
- 10 to 100 times the storage capacity of conventional solid state memory

Magnetic Race Track Memory
S. Parkin (IBM), US patents
Magnetic anisotropy at a surface

- Free atomic spin is rotationally invariant: all spin orientations are degenerate.
- Loss of rotational symmetry breaks degeneracy of spin orientations.

\[ H = -g\mu_B \mathbf{B} \cdot \mathbf{S} + DS_z^2 \]

Magnetic field dependence varies with angle of magnetic field.
The energy that is required to change the direction of a single spin on CuN measured.

- Large single-atom magnetic anisotropy for iron of about 6 meV.
- About 50x weaker anisotropy for manganese on same surface.

Spin excitation spectroscopy reveals spin energy levels, including their magnetic field dependence.

DFT calculations elucidate surface structure and leads to same total spin as experiment.

**GOAL**: engineer very large magnetic anisotropy to demonstrate data storage.
The Future of Memory?

• Phase-change memory – low cost because >1 bit / 4F^2
  ~2013?

• Racetrack memory
  ~2018? – a 3-D nano-warehouse for data

• Atomic memory – “there’s a lot of room at the bottom…”
  ~2030?
Thank you!