Advanced Memory Technology
- #1 Factor for Energy Efficient System -

2012. 09

Kenny Han

Samsung Semiconductor Inc.
This presentation is intended to provide information concerning supercomputer and memory industry. We do our best to make sure that information presented is accurate and fully up-to-date. However, the presentation may be subject to technical inaccuracies, information that is not up-to-date or typographical errors. As a consequence, Samsung does not in any way guarantee the accuracy or completeness of information provided on this presentation. Samsung reserves the right to make improvements, corrections and/or changes to this presentation at any time.

The information in this presentation or accompanying oral statements may include forward-looking statements. These forward-looking statements include all matters that are not historical facts, statements regarding the Samsung Electronics' intentions, beliefs or current expectations concerning, among other things, market prospects, growth, strategies, and the industry in which Samsung operates. By their nature, forward-looking statements involve risks and uncertainties, because they relate to events and depend on circumstances that may or may not occur in the future. Samsung cautions you that forward looking statements are not guarantees of future performance and that the actual developments of Samsung, the market, or industry in which Samsung operates may differ materially from those made or suggested by the forward-looking statements contained in this presentation or in the accompanying oral statements. In addition, even if the information contained herein or the oral statements are shown to be accurate, those developments may not be indicative developments in future periods.
1. Introduction
2. DRAM Requirements in Exa-Scale Computing
3. DRAM Technology Evolution & Challenges
4. DRAM Solution Consideration
5. SSD for Storage Solution
6. New Memory Technology: STT-MRAM
7. Summary
Why Exa-Scale Computing?

- Data explosion by social media, on-line games, cloud computing, ...
- Over 1 ZETA bytes in 2010 and keep rapidly growing

![Graph showing created data (Exabyte) and supercomputer FLOPS over years.](image)

*Log Scale*

[Created Data (Exabyte)](image) vs [No.1 Super Computer PFLOPS](image)

[Source: IDC(2011) & Top500.org(2011) & Estimated by extrapolation]
Memory Wall

Widening of the gap between CPU and Memory performance

- Memory becomes more important than ever
- Requires larger capacity, higher performance, and better power efficiency

<table>
<thead>
<tr>
<th></th>
<th>Exa-Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory Capacity (System)</td>
<td>~10PB</td>
</tr>
<tr>
<td>Memory BW (Node)</td>
<td>0.5~1TB/s</td>
</tr>
<tr>
<td>Power (System)</td>
<td>~20MW</td>
</tr>
</tbody>
</table>

[Effective CPU cycle time, CPU cycle time, SRAM access time, DRAM access time, Flash SSD access time, Disk seek time]

[Effective CPU cycle time, CPU cycle time, SRAM access time, DRAM access time, Flash SSD access time, Disk seek time]

[Carnegie Mellon Univ. & Samsung Estimation]
1. Introduction

2. DRAM Requirements in Exa-Scale Computing

3. DRAM Technology Evolution & Challenges

4. DRAM Solution Consideration

5. SSD for Storage Solution

6. New Memory Technology: STT-MRAM

7. Summary
Capacity Requirements in Exa-Scale Computing

- Requires more than 70x memory capacity

Memory Capacity Requirements

[Source: “Memory systems for PetaFlop to ExaFlop class machines” by IBM, 2007 & 2010]
Bandwidth Requirements in Exa-Scale Computing

Requires more than 100x bandwidth per node

Memory Bandwidth Requirements

Now

2018

10~20GB/s

400~600Mbps

7.5x

(~100GB/s)

3.7x

(~1.6Gbps)

100x

(~1.4TB/s)

12.5x

(~5.3Gbps)

Peta-flops

20Peta-flops

Exa-flops

[Source: “Memory systems for PetaFlop to ExaFlop class machines” by IBM, 2007 & 2010]
Power Efficiency Requirements in Exa-Scale Computing

Requires at least doubled memory power efficiency
- Larger capacity and higher bandwidth require extreme power efficiency

---

**Power Efficiency**

- Now: 1x
- 2018: \(~0.3x\) for W/Gbps

---

**Memory Power Projection**

- Now: 37x
- 2018: 16x

---

[Source: "Memory systems for PetaFlop to ExaFlop class machines" by IBM, 2007 & 2010]
Contents

1. Introduction
2. DRAM Requirements in Exa-Scale Computing
3. DRAM Technology Evolution & Challenges
4. DRAM Solution Consideration
5. SSD for Storage Solution
6. New Memory Technology : STT-MRAM
7. Summary
Larger capacity mono die introduction is getting delayed

- Technology difficulty & tremendous investment
- Disruptive technology is necessary to increase module capacity more

Subject to cost/energy efficiency, scaling, ...

* Highest Module Capacity with Mono-component
Feeding data with enough BW will be the 1st Challenge for exa-scale Computing.

- **Limitation by memory I/F & channel environment**
- **Limitation by physical channel environment & stacking technology**

*Highest speed in that year*
Channel Environment Challenge

Limitation of physical channel environment

• Validity of multi-drop bus architecture beyond DDR4 to support > 25.6GBps/Ch.?

New approach for memory sub-system

• Differential signaling or serial link?
• New bus architecture like point-to-point, daisy-chained, ... 
• New technology like optical I/O

<table>
<thead>
<tr>
<th></th>
<th>DDR4</th>
<th>Post DDR4*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed/pin</td>
<td>1.6~3.2Gbps</td>
<td>~6.4Gbps</td>
</tr>
<tr>
<td>BW/Ch.</td>
<td>~25.6GB/s</td>
<td>~51.2GB/s</td>
</tr>
</tbody>
</table>

* Expectation
**Memory Power Efficiency Trend**

- Aggressive process shrink provides better power efficiency and performance.
- 20nm class 4Gb 1.35V can offer the most efficient power usage.

### Memory Power Consumption as measured in 96GB server

<table>
<thead>
<tr>
<th>Technology</th>
<th>Interface</th>
<th>Density</th>
<th>Voltage</th>
</tr>
</thead>
<tbody>
<tr>
<td>50nm class</td>
<td>DDR3</td>
<td>2Gb</td>
<td>1.5V</td>
</tr>
<tr>
<td>40nm class</td>
<td>DDR3</td>
<td>2Gb</td>
<td>1.5V</td>
</tr>
<tr>
<td>30nm class</td>
<td>DDR3</td>
<td>2Gb</td>
<td>1.35V</td>
</tr>
<tr>
<td>20nm class</td>
<td>DDR3</td>
<td>4Gb</td>
<td>1.35V</td>
</tr>
<tr>
<td>20nm class</td>
<td>DDR3</td>
<td>4Gb</td>
<td>1.25V*</td>
</tr>
</tbody>
</table>

- 65.3W *34% Savings*
- 42.8W *21% Savings*
- 33.6W *35% Savings*
- 21.8W *67% Savings*
- 20.3W

*Optional Solution*

*Considered with an 8 hours active and 16 hours idle status in server.*

Source: Samsung Lab.
Efficiency improvement by process shrink only is limited
- Additional technology aids to reduce the power consumption more
Technology Scaling Challenges

**DRAM is not free any more**

- Scaling approaches a theoretical limit
- Technology difficulty & large investment
  → Very difficult to achieve economical scaling

---

Cost/bit

Investment/WF

Transition Period

40nm class  30nm class  20nm class  10nm class  Sub 10nm class

2011  2015  2018

Requires Close-Collaboration between All Parties: CPU, Memory, S/W, etc
1. Introduction
2. DRAM Requirements in Exa-Scale Computing
3. DRAM Technology Evolution & Challenges
4. DRAM Solution Consideration
5. SSD for Storage Solution
6. New Memory Technology : STT-MRAM
7. Summary
**TSV (Through-Silicon Via)**

- **TSV** offers less power consumption and higher performance by hiding electrical loading
  - Still concern on cost, even if matured, *intrinsic cost overhead exist*...

### Power Comparison

<table>
<thead>
<tr>
<th></th>
<th>TSV</th>
<th>Conventional Stacking</th>
</tr>
</thead>
</table>
| **Pros**         | • Short Interconnection (< ~50um)  
                    • Lower Profile  
                    • More # of Interconnects (>1000ea)  | Low Cost & Matured Technology |
| **Cons**         | **High Cost**                 |                                   |
|                  | • Long Loop Wires  
                    • Higher Profiles  
                    • Limitations in # of Interconnects  
                    • Overhang         |                                   |

**Merits**
- 3DS-TSV consumes 17% lower power

*Measured by 32GB RDIMM @ 2DPC*
Optical Interface

- Mainly has been utilized in long distant communication

**Interface Power (I/O & Termination) getting dominant**

- Memory sub-system is relatively short channel environment
- Potential to get ~60% better power efficiency over DDR3 interface
- Opportunity to expand # of slots and support high pin speed
- How to implement power efficient and low cost optical I/O solution?

---

![I/F power getting dominant](image.png)

**Optical I/O: lower Power per B/W**

- [Graph showing power cost vs. bandwidth for different technologies]

* Source: Samsung
Controller Offloading

**Potential advantages by additional functionalities in logic die**

- Distributed small scale computing
- Reduced controller complexity & increased performance
- Better system power efficiency with reduced data traffic
- Additional logic to enhance device reliability and DRAM scaling extension
- Supporting heterogeneous memories → DDR3, DDR4, PRAM, Flash, MRAM etc.
Memory Sub-System Candidate

Important to figure out the memory sub-system satisfying Capacity, Performance, and Power efficiency requirements of HPC

Memory Cube Module

- Ch.0
- Ch.1

Optical I/F

Multi-drop Channel

Large Capacity

Large Capacity & Power Efficiency

High BW & Power Efficiency

Additional Layer with Higher BW, small capacity DRAM

L1

L2

Memory Cube

DDR4
1. Introduction

2. DRAM Requirements in Exa-Scale Computing

3. DRAM Technology Evolution & Challenges

4. DRAM Solution Consideration

5. **SSD for Storage Solution**

6. New Memory Technology : STT-MRAM

7. Summary
Huge latency gap between memory and disk drive

- Only 1.3X mechanical access improvement in HDD for 13 years
- Flash storage is a good gap filler to minimize the distance

"I/O Memory Tier"
(Typically 50us ~ 300us Latency)

Flash storage is a good candidate to fill the gap
SSD provides extremely higher performance than disk drive

- Unlocking true NAND Flash potential with host interface enhancement
- >20x Latency, >100x IOPs, and >3x Sequential Performance
Power Saving with SSD

Enhanced energy efficiency and smaller footprint through SSD

- Higher single device performance reduces number of drives
- A case study shows 73% power reduction with a half number of drives

![Graph showing energy savings.

Aggregated SAS: 15K SAS

HDD only: 100% 15K RPM 600GB SAS 167ea.

HDDs+SSDs: 5% 200GB SSD 25ea., 95% 7.4K RPM 2TB SATA 48ea.

Tired Storage: 7.5K SATA

73% reduction/system
Smaller footprint (~ ½ number of drives)

Source: EMC, 2011
Storage Capacity: 100TB

Case Study

[Graph showing energy consumption comparison between HDD only and HDDs+SSDs.]
NAND Flash Innate Characteristics

NAND flash management technique is different

- **NAND characteristics**
  - No overwrite
  - Page operation but block erase

- **NAND suffers from**
  - P/E cycle
  - Data retention

High Performance & Reliable SSD

- **Flash Aware Host System**
  (Trim, Log-Structured File System, Page alignment shaping)

- **Flash Abstraction Layer**
  (Wear Leveling, Bad Block Management, Garbage Collection)

- **Flash Physical Layer**
  (Error Correction, LDPC, Scrambling & High Endurance Features)

Sophisticated NAND Flash Management Technology

VS.

Align with your imagination
Managing Endurance

Facing challenges to maintain P/E cycles with process shrink

- To extend SSD life time, it needs sophisticated HW & SW technologies
  - Continued Evolution in Controller Technology
  - Adaptive Management & Tuned NAND Flash
  - Application Awareness

NAND Flash Endurance

- SLC Endurance
- MLC Endurance

Years

SSD Requirement

Shrink Rate Slows Down
Reliability Degrades
Performance Deteriorates
Active collaboration is the key for SSD value capture in HPC

- SSD characteristics is too variant to be a generic solution for all storage needs
- Engineered systems allows to accelerate the benefits of flash while avoiding the pitfalls
Contents

1. Introduction
2. DRAM Requirements in Exa-Scale Computing
3. DRAM Technology Evolution & Challenges
4. DRAM Solution Consideration
5. SSD for Storage Solution
6. New Memory Technology : STT-MRAM
7. Summary
New Memory Technology: STT-MRAM

Non-volatility and comparable to DRAM
- Highly desired solution, but technology maturity?

STT-MRAM Cell Structure
How to Capture Non-Volatility Value?

Collaboration to recognize as a value added device

- Maximize values and advantages of non-volatility in STT-MRAM
- Build up eco-system: H/W, S/W including OS, ...
  - Software overhead minimization
  - New architecture combining memory and storage
  - New application with non-volatile buffer
- Additional value creation will dilute initial large investment

[Many hurdles]
- Initial large investment
- Technology difficulty
- Process maturity
- Etc...

---

Simple Replacement of DRAM?
1. Introduction
2. DRAM Requirements in Exa-Scale Computing
3. DRAM Technology Evolution & Challenges
4. DRAM Solution Consideration
5. SSD for Storage Solution
6. New Memory Technology : STT-MRAM
7. Summary
Exa-Scale computing requires large capacity, high bandwidth, and power efficient memory sub-system.

DRAM technology is getting difficult and expensive.

A revolutionary and innovative memory solution is mandatory
- Continued scaling, TSV, Optical IO, Controller offloading, ...

SSD is a good candidate for efficient storage system
- Very high performance and power efficient device
- Different characteristics require different optimization

Very high potential in STT-MRAM, but technology is not yet

Call to action for active and strong collaboration between all parties
Thank you