Low-Power, High-Bandwidth, and Ultra-Small Memory Module Design

Qawi I. Harvard, PhD, and R. Jacob Baker, PhD
Department of Electrical and Computer Engineering
Abstract

The main memory subsystem has become inefficient. Sustaining performance gains has power consumption, capacity, and cost moving in the wrong direction. This talk proposes novel module, DRAM, and interconnect architectures in an attempt to alleviate these trends. The proposed architectures utilize inexpensive innovations, including interconnect and packaging, to substantially reduce the power, and increase the capacity and bandwidth of the main memory system. A low cost advanced packaging technology is used to propose an 8 die and 32-die memory module. The 32-die memory module measures less than 2 cm³. The size and packaging technique allow the memory module to consume less power than conventional module designs. A 4 Gb DRAM architecture utilizing 64 data pins is proposed. The DRAM architecture is inline with ITRS roadmaps and can consume 50% less power while increasing bandwidth by 100%. The large number of data pins are supported by a low power capacitive-coupled interconnect. The receivers developed for the capacitive interface were fabricated in 0.5 µm and 65 nm CMOS technologies. The 0.5 µm design operated at 200 Mbps, used a coupling capacitor of 100 fF, and consumed less than 3 pJ/bit of energy. The 65 nm design operated at 4 Gbps, used a coupling capacitor of 15 fF, and consumed less than 15 fJ/bit.
Mobile Platform

- Motorola Atrix (Front)
  - Linear Power Amplifier
  - RF Power Management
  - DC Power Management
  - Hard Drive
  - Accelerometer

< 63.5 mm

< 117.75 mm
Mobile Platform

- Motorola Atrix (Back)

- Memory (DSP)
- Memory & CPU
- HSPA+ DSP
- 802.11n & Bluetooth
- Compass
Server Platform

- Intel Server Board S5502UR

- Memory Slots
Organization

- Main Memory Limitations
- Nano-Module
- Wide I/O DRAM Architecture
- High Bandwidth Interconnect
- Conclusions
Main Memory Limitations

- Datacenter sparsity masked power limitations
  - Power trend: Energy consumption doubled every 5 years
- Server power
  - ~50 W in 2000
  - ~250 W in 2008
- Server power breakdown
  - CPU: 37%, Memory: 17%
  - Trend is Memory power > CPU power
- Main memory power
  - More die per module
  - Less modules per channel
  - Higher bandwidth
Main Memory Limitations

- CPU: 37%
- Memory: 17%
- PCI Slots: 23%
- Disk: 6%
- Fan: 5%
- Motherboard: 12%
Main Memory Limitations

![Graph showing global spending on new server and power and cooling over years 1996 to 2010. The graph uses blue for new server spending and red for power and cooling spending. The y-axis represents global spending in billions ($), and the x-axis represents years from 1996 to 2010.]
Main Memory Limitations

- CPU power wall
  - Voltage scaling reached its limit
  - Multiple cores supplement performance gains
  - No “multi-core” for DRAM
- DRAM voltage scaling reaching its limit
  - Current rate increase > voltage reduction rate
  - Power increasing
- DRAM pre-fetch
  - Memory core operates at slower frequency
  - High power I/O devices and data-path
Main Memory Limitations
Main Memory Limitations

- DRAM inefficiencies increase cost and power
  - Processor cache increasing
  - Intel Nehalem processor
  - DRAM would need to have L3 BW and latency
  - “…create the illusion of a large memory that we can access as fast as a very small memory.” – Patterson & Hennessy

<table>
<thead>
<tr>
<th></th>
<th>Local</th>
<th>L1</th>
<th>L2</th>
<th>L3</th>
<th>RAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read BW</td>
<td></td>
<td>45.6</td>
<td>31.1</td>
<td>26.2</td>
<td>10.1</td>
</tr>
<tr>
<td>Write BW</td>
<td></td>
<td>45.6</td>
<td>28.8</td>
<td>19.9</td>
<td>8.4</td>
</tr>
<tr>
<td>Latency</td>
<td>(ns)</td>
<td>1.3 (4)</td>
<td>3.4 (10)</td>
<td>13.0 (38)</td>
<td>65.1 (191)</td>
</tr>
</tbody>
</table>
Main Memory Limitations

- DRAM efficiencies increase performance
- Capacity versus Performance

- Capacity costs power
  - Multiple memory channels
  - Each additional module increases power
Main Memory Limitations

- Bandwidth versus performance

- Bandwidth costs power
  - Buffer on board
  - Multiple channels
Main Memory Limitations

- DRAM inefficiencies in practice
- Typical video/web server motherboard
  - ✔️ 20+ layer PCB
  - ✔️ 6 memory channels
- RDIMM
  - ✔️ 10+ layer PCB
  - ✔️ Maximum comp. count
Main Memory Limitations

12 RDIMM

- **Termination**
  - 36 components per DIMM
  - 8 I/O per component
  - 2.7 W of termination power for a read/write per module
  - 32.4 W total termination power

- **Wordline firing**
  - 100 ns activation rate
  - 8126 page size
  - 200 fF per bitline
  - 11.2 W total bitline sense amplifier power

Sustaining performance gains through capacity and bandwidth increases power and cost – innovation required.
Nano-Module

- **Goals**
  - Purpose was to move labs into prototype generation
  - Required low cost, high bandwidth, and low power memory solution that can be used with capacitive coupled interconnects in advanced server architectures

- **Module component count trends required a new approach**

- **Nano-module proposed**
  - Low cost advanced packaging technology
  - Off-the-shelf memory components

- **Results can be leveraged**
  - NAND
  - Mobile
Nano-Module

- Literature review of high capacity memory stacks
- 1990’s
  - Multichip Modules
    - Realized planar space limitations
  - Val & Lemione
  - Irvine Sensors
- Solutions proposed in research
  - No industry due to memory hierarchy effectiveness
Nano-Module

- Memory stack technology gaining new attention
- 2010
  - Samsung quad die with TSV
  - 80 µm pitch, 30 µm diameter, 300 TSV
  - $R_{TSV} = 5 \, \Omega$, $C_{TSV} = 300 \, fF$
- Pros:
  - Lower power, higher bandwidth
- Cons:
  - Cost, integration

![Diagram of Nano-Module with TSV and chips arrangement]
Nano-Module

- Literature review revealed novel solutions
- Slant the die!
- Applicable to capacitive-coupled interconnects
Nano-Module

- Not the first to try it:
Nano-Module

- **Controlled Impedance**
  - ✓ All Signals 50 Ω controlled impedance
  - ✓ DQS and CLK 120 Ω differential impedance

- **Trace Length Matching**
  - ✓ All Data matched to worst case
  - ✓ All CLK matched to worst case
  - ✓ All Address/Command matched to worst case

\[
Z_0 = \frac{87}{\sqrt{\varepsilon_r + 1.41}} \ln\left(\frac{5.98H}{0.8W + T}\right)
\]

Microstrip
Size calculations

\[
\text{height} = t_{\text{sub thickness}} + t_{\text{connection}} + \frac{t_{\text{die thickness}}}{\cos \alpha} + \left( t_{\text{die width}} - t_{\text{die pad to edge}} \right) \sin \alpha
\]

\[
\text{width} = (\text{no. signals} - 1) t_{\text{con pitch}} + t_{\text{con diameter}} + 2 \left( t_{\text{die pad to edge}} + t_{w_{b_1}} + t_{w_{b_2}} \right)
\]

\[
\text{length} = 2 \left( t_{w_{b_1}} + t_{w_{b_2}} \right) + \sin \alpha \cdot t_{\text{die thickness}} + \frac{\left( \# \text{ die} - 1 \right) t_{\text{die thickness}}}{\sin \alpha} + \cos \alpha \cdot t_{\text{die width}}
\]
Nano-Module

- Thermal option
  - Thermal conductivity
    - Silicon, Metals >> Mold Compound
    - Hot spots
    - Temperature gradient
Nano-Module

- Thermal option
  - Heat plate
Wide I/O DRAM Architecture

- **4 Gb DRAM**
  - Meets 2012 ITRS predictions
  - Developed at Boise State
- **Edge aligned pads**
- **Page size reduction**
- **Low cost process**
  - < 4 levels of metal
  - No impact to die size
  - No impact to array efficiency
- **Move to 64 data pins**
  - Report challenges
  - Propose innovations

<table>
<thead>
<tr>
<th>Chip Size = 71.4 mm²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Array Efficiency = 57.7%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>256M Array</th>
<th>Row</th>
<th>256M Array</th>
<th>Row</th>
<th>256M Array</th>
<th>Row</th>
<th>256M Array</th>
</tr>
</thead>
<tbody>
<tr>
<td>Column</td>
<td></td>
<td>Column</td>
<td></td>
<td>Column</td>
<td></td>
<td>Column</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>256M Array</th>
<th>Row</th>
<th>256M Array</th>
<th>Row</th>
<th>256M Array</th>
<th>Row</th>
<th>256M Array</th>
</tr>
</thead>
<tbody>
<tr>
<td>Column</td>
<td></td>
<td>Column</td>
<td></td>
<td>Column</td>
<td></td>
<td>Column</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>256M Array</th>
<th>Row</th>
<th>256M Array</th>
<th>Row</th>
<th>256M Array</th>
<th>Row</th>
<th>256M Array</th>
</tr>
</thead>
<tbody>
<tr>
<td>Column</td>
<td></td>
<td>Column</td>
<td></td>
<td>Column</td>
<td></td>
<td>Column</td>
</tr>
</tbody>
</table>

College of Engineering
Wide I/O DRAM Architecture

- 4 Gb Edge DRAM
  - Centralized Row and Column
  - Smaller die
  - Higher efficiency
  - < 4 levels of metal

Chip Size = 68.88 mm²
Array Efficiency = 59.9%

12.3 mm

512M Bank    512M Bank
512M Bank    512M Bank
COLUMNS     COLUMNS
512M Bank    512M Bank
512M Bank    512M Bank
Edge I/O Interface
Wide I/O DRAM Architecture

- **Challenges**
  - ✔ Number of metal layers
  - ✔ Global data routing
  - ✔ Local data routing

- **Proposals**
  - ✔ Split bank structure
  - ✔ Data-path design
  - ✔ Through bitline routing
  - ✔ SLICE architecture
  - ✔ Capacitive-coupled I/O
High Bandwidth Interconnect

- Capacitive-coupling
  - Increased bandwidth
    - Reduced ESD capacitance
    - Smaller I/O channel = more I/O
    - Removal of inductive channel
  - Low power
    - Reduced ESD capacitance
    - Low power Tx & Rx
  - Low cost
    - Simple
  - Alignment required

- Literature review
  - Revealed inefficiencies and lack of application
High Bandwidth Interconnect

- Proposed receiver design
  - Extreme low power
  - $\sim1$ gate delay latency
  - ‘DC’ transmission
  - RTZ $\rightarrow$ NRZ
High Bandwidth Interconnect

- 0.5 µm CMOS design (proof of concept)
  - 5.0 V process
  - 50 fF poly-poly capacitor
  - 200 Mbps
  - 3 – 8 pJ/bit
  - 325 Gb/mm²
High Bandwidth Interconnect

- Chip micrograph
  - 1.5 mm x 1.5 mm
  - 9 structures

- Experimental results
  - Operate at $V_{TX} = 2.0$ V
  - 3 pJ/bit at 200 Mbps
High Bandwidth Interconnect

- 65 nm CMOS design (proof of scalability)
  - 1.2 V process
  - 15 fF metal-metal capacitor
  - 4 Gbps
  - 17 µm²
  - 227 Tbps/mm²
High Bandwidth Interconnect

- **Die micrograph**
  - 2 mm x 2 mm

- **Experimental results**
  - 2 Gbps @ 0.9V
  - 50 fF coupling capacitor

1.0 Gbps – 1.2 V

2.0 Gbps – 0.9 V
## High Bandwidth Interconnect

<table>
<thead>
<tr>
<th>Work</th>
<th>Process</th>
<th>Supply</th>
<th>Data Rate</th>
<th>Coupling</th>
<th>Gbps/mm²</th>
<th>Energy</th>
<th>Requires CLK?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kanda, 2003</td>
<td>0.35 µm</td>
<td>3.3 V</td>
<td>1.27 Gbps</td>
<td>~10 fF</td>
<td>2117</td>
<td>2.4 pJ/bit</td>
<td>Yes</td>
</tr>
<tr>
<td>Wilson, 2007</td>
<td>0.18 µm</td>
<td>1.8 V</td>
<td>3 Gbps</td>
<td>150 fF</td>
<td>555</td>
<td>5 pJ/bit</td>
<td>No</td>
</tr>
<tr>
<td>Fazzi, 2008</td>
<td>0.13 µm</td>
<td>1.2 V</td>
<td>1.23 Gbps</td>
<td>~10 fF</td>
<td>19,219</td>
<td>0.14 pJ/bit</td>
<td>Yes</td>
</tr>
<tr>
<td>Kim, 2009</td>
<td>0.18 µm</td>
<td>1.8 V</td>
<td>2 Gbps</td>
<td>600 fF</td>
<td>690</td>
<td>0.8 pJ/bit</td>
<td>Yes</td>
</tr>
<tr>
<td>This work</td>
<td>0.5 µm</td>
<td>5.0 V</td>
<td>200 Mbps</td>
<td>50 fF</td>
<td>325</td>
<td>8 pJ/bit</td>
<td>No</td>
</tr>
<tr>
<td>This work</td>
<td>65 nm</td>
<td>1.2 V</td>
<td>4 Gbps</td>
<td>15 fF</td>
<td>226,757</td>
<td>0.015 pJ/bit</td>
<td>No</td>
</tr>
</tbody>
</table>


Conclusions

- **Nano-Module**
  - Developed a new research direction for industry research labs
  - Developed initial motivation
  - Developed initial prototype

- **DRAM Architecture**
  - Demonstrated benefits of wide I/O topologies
  - Proposed several low power innovations
  - Provided application for novel interconnect technologies

- **Capacitive-Coupled Receiver**
  - Demonstrated low power receiver designs
  - Achieved 2 Gbps at < 15 fJ/bit in 65 nm
Questions
Appendix - PLL

- 65 nm test chip
  - PLL
  - PRBS generator
Appendix - PLL

- PLL
  - Phase detector
Appendix - PLL

- Charge pump

\[ A_{PD} = \frac{I_{PUMP}}{2\pi} \]

\[ A_F = \frac{1 + sR_1C_1}{sC_1} \]
Appendix - PLL

- Voltage controlled oscillator

\[ A_{VCO} = 2\pi \cdot \frac{f_{MAX} - f_{MIN}}{V_{MAX} - V_{MIN}} \]
Appendix - PLL

- Over damped

\[
\phi_{\text{CLK}_{\text{out}}} = V_{\text{IN}_{\text{VIN}}}, \frac{A_{\text{VCO}}}{s}
\]

\[
\phi_{\text{CLK}_{\text{d}}} = \frac{1}{N}, \quad \phi_{\text{CLK}_{\text{out}}} = \beta \cdot \phi_{\text{CLK}_{\text{d}}}
\]

\[
A_{\text{OL}} = A_{PD}A_{F}A_{\text{VCO}}
\]

\[
H(s) = \frac{\phi_{\text{CLK}_{\text{in}}}}{\phi_{\text{CLK}_{\text{out}}}} = \frac{A_{PD}A_{F}A_{\text{VCO}}}{s} = \frac{A_{PD}A_{F}A_{\text{VCO}}}{s + \beta A_{PD}A_{F}A_{\text{VCO}}}
\]

\[
A_{F} = \frac{1 + sR_{C_{1}}}{sC_{1}}
\]

\[
H(s) = \frac{\phi_{\text{CLK}_{\text{in}}}}{\phi_{\text{CLK}_{\text{out}}}} = \frac{A_{PD}A_{\text{VCO}}}{s^{2} + A_{PD}A_{\text{VCO}} \left( \frac{1 + sR_{C_{1}}}{C_{1}} \right)}
\]

\[
\omega_{n} = \sqrt{\frac{A_{PD}A_{\text{VCO}}}{NC_{1}}}
\]

\[
\zeta = \frac{\omega_{n}}{2 R_{C_{1}}}
\]
Appendix - PLL

- PLL at lock
Appendix - PLL

- PLL layout
Appendix - PLL

- PRBS generator

![PRBS generator diagram and eye diagram](image-url)
Appendix - PCB

- PCB test board
Appendix – Dead Bug
Appendix – 65 nm Chip
References


[3] Uksong Kang; Hoe-Ju Chung; Seongmoo Heo; Duk-Ha Park; Hoon Lee; Jin Ho Kim; Soon-Hong Ahn; Soo-Ho Cha; Jaesung Ahn; DukMin Kwon; Jae-Wook Lee; Han-Sung Joo; Woo-Seop Kim; Dong Hyeon Jang; Nam Seog Kim; Jung-Hwan Choi; Tae-Gyeong Chung; Jei-Hwan Yoo; Joo Sun Choi; Changhyun Kim; Young-Hyun Jun; , "8 Gb 3-D DDR3 DRAM Using Through-Silicon-Via Technology," *Solid-State Circuits, IEEE Journal of*, vol.45, no.1, pp.111-119, Jan. 2010


References