Nonvolatile Erasable NAND Flash Memories

Raul Adrian Cernea
SanDisk Corporation, Milpitas, California
Outline

• NAND Flash and memory cards
• Basic operations and multi-level
• Performance trends and ABL
• 3X MLC program-throughput
• 3 bits per cell
• Summary and conclusions
Mobile/CE/PC Driving NAND Future
Memory cards

Every card:

⇒ One controller
⇒ At least one memory chip
Erase and program
Cross-section along the Word Line

WORD LINE

FLOATING GATES

NON-CONDUCTIVE NAND CHAIN

CONDUCTIVE NAND CHAIN
Cross-section along the Bit Line

- WL
- FG
- CONDUCTIVE NAND CHAIN
- NON-CONDUCTIVE NAND CHAIN
Program and verify

![Diagram showing Program and Verify with SLC and MLC](image-url)
SLC and MLC

NUMBER OF CELLS

SINGLE LEVEL CELL

MULTI LEVEL CELLS

TWO BITS PER CELL

THREE BITS PER CELL

FOUR BITS PER CELL
Performance trends and All Bit Line Architecture
Program-throughput trends for NAND memories

<table>
<thead>
<tr>
<th>Year</th>
<th>SLC</th>
<th>MLC</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>10MB/s</td>
<td>16.7MB/s</td>
</tr>
<tr>
<td>2004</td>
<td>20MB/s</td>
<td>2.7MB/s</td>
</tr>
<tr>
<td>2005</td>
<td>30MB/s</td>
<td>2.1MB/s</td>
</tr>
<tr>
<td>2006</td>
<td>40MB/s</td>
<td>7.7MB/s</td>
</tr>
<tr>
<td>2007</td>
<td>50MB/s</td>
<td>11MB/s</td>
</tr>
<tr>
<td>2008</td>
<td>60MB/s</td>
<td>18.4MB/s</td>
</tr>
</tbody>
</table>

**CURRENT 16Gb ABL**
- ISSCC 2008

**ABL**
- WITHOUT HIGH SPEED I/O INTERFACE

**ABL MLC**
- FULL SEQUENCE

**ABL MLC**
- CONVENTIONAL
The 16Gb chip
Cell access

Y DECODER: ACCESS TO ALL BIT LINES

X DECODER: SELECTED WORD LINE

X DECODER: SELECTED BLOCK

X DECODER: UNSELECTED BLOCK(S)

SENSE AMPLIFIER

SENSE AMPLIFIER

SENSE AMPLIFIER

SENSE AMPLIFIER

32 NAND CELLS
All Bit Line memory versus “Conventional”

- Simultaneous access of two word lines of the same size
- ABL: double Column Logic, two sided
- 8 pages of 4KB parallel programming (4 pages conventional)
- Maximum MLC program-throughput 34MB/s (10MB/s conventional)
Maximum parallelism

• Conventional is an Even / Odd architecture
  — One Sense Amplifier handling two Bit Lines
  — Every other Bit Line shielded during sensing

• Full READ and PROGRAM capabilities for ABL
  — No shielding necessary (as in current mainstream NAND architecture)
Conventional column architecture

- Two Bit Lines for every read-program unit
- All similar column blocks
ABL: three level hierarchical architecture

<table>
<thead>
<tr>
<th>CONTROL LOGIC</th>
<th>CONTROL LOGIC</th>
<th>CONTROL LOGIC</th>
<th>CONTROL LOGIC</th>
<th>COMMON BUS</th>
<th>N</th>
<th>CONTROL LOGIC</th>
<th>CONTROL LOGIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLOCK SELECTION</td>
<td>POINTER</td>
<td>INPUT / OUTPUT</td>
<td>LOCAL DRIVERS</td>
<td>LOCAL DRIVERS</td>
<td>LOCAL DRIVERS</td>
<td>LOCAL DRIVERS</td>
<td>LOCAL DRIVERS</td>
</tr>
</tbody>
</table>

- One SA per BL means *double column circuitry*
  - Area reduction is a must
- Hierarchical architecture: less repeating circuits
  - Only one Sequential Logic Unit and one Control Logic per group
  - Only one Block Selection, one Pointer and one I/O per block
- Saved area allows for local drivers
3X MLC Program-Throughput
MLC programming time

- There is a step up in programming time when one compares the lower page to the upper page (one state versus two states)
- The lower and upper pages together (three states) are the slowest to be programmed
• In ABL architecture, programming four pages in parallel is twice as fast as Conventional
• MLC programming in Full-Sequence is about three times faster than Conventional
Yupin Effect

• Additional negative voltage in a two-step sequence for Floating Gates programmed in step TWO (Yupin Effect*)

Ref.: SanDisk US Patent 5,867,429 (Feb. 2 1999)

• Minimized effect in one-step programming (Full-Sequence)
ABL: faster verify with current sensing

- Pre-charge (by Bit Line driver): non-linear, fast (both)
- Discharge (by NAND cell): linear, slow (Conventional)
- Faster verify operation in ABL mode (pre-charge only)
Two step I/O redundancy access

- Regular clock for redundancy bytes at Data In (or Data Out)
  - **STEP 1:** Full data stream into the regular buffer
  - **STEP 2:** Data transfer between regular buffer and redundancy zone
- Reversed process for Data Out
- Minimal time penalty for the hidden data transfer
ABL for lower energy

- Even / Odd sensing in Conventional architecture
  - Double sensing operations per Word Line
  - Strong coupling between adjacent Bit Lines plus parasitic to Ground
- All Bit Line sensing in one operation
  - Bit Line coupling to ground only (20% of total parasitic)
  - Lower Bit Line voltages
- At least 10 times less charge wasted when using ABL
• A full Word Line distribution of random data, programmed in ABL, Full-Sequence at a 34MB/s program-throughput
Three bits per cell
Memory Density Trend

- This work (NAND X3)
- Samsung (NAND)
- Toshiba/SanDisk (NAND)
- SanDisk/Toshiba D2
- Hitachi (AND)
- Toshiba/SanDisk (NAND)

Density (Mb/mm²) vs. Year

- 2000
- 2002
- 2004
- 2006
- 2008
- 2010

37% Mb/mm²
# Chip architecture comparison

<table>
<thead>
<tr>
<th></th>
<th>56nm 8G D2</th>
<th>56nm 16G D2</th>
<th>56nm 16G D3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chip size</strong></td>
<td>99mm²</td>
<td>182mm²</td>
<td>142.5mm² (this work) (22% saving)</td>
</tr>
<tr>
<td><strong>Density</strong></td>
<td>8Gb (80.8Mb/mm²)</td>
<td>16Gb (82Mb/mm²)</td>
<td>16Gb (112Mb/mm²)</td>
</tr>
<tr>
<td><strong>Architecture</strong></td>
<td>4Gb / plane</td>
<td>8Gb / plane</td>
<td>8Gb / plane</td>
</tr>
<tr>
<td><strong>Bits/Cell</strong></td>
<td>2bits/cell</td>
<td>2bits/cell</td>
<td>3bits/cell</td>
</tr>
<tr>
<td><strong>Program/Sense</strong></td>
<td>Even or Odd</td>
<td>ABL</td>
<td>ABL</td>
</tr>
</tbody>
</table>

(22% saving)
3 bit per cell programming

- 3 pages (Lower/Middle/Upper) on each Word Line
- Each page can be treated as an independent page
- The pages have to be programmed sequentially
3 bit per cell programming and verify

- 7 verify operations after each programming pulse
- Conversion from “page by page” to three page programming for speed improvement
- Program-throughput comparable to conventional (non-ABL 2 bit per cell)
7 state distribution
Summary and Conclusions
## Summary, 2 bit per cell

<table>
<thead>
<tr>
<th>READ THROUGHPUT BY 8 I/O</th>
<th>50MB/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABL SLC PROGRAMMING (WITHOUT HIGH SPEED I/O INTERFACE)</td>
<td>60MB/s</td>
</tr>
<tr>
<td>ABL MLC PROGRAMMING NO FULL-SEQUENCE</td>
<td>24MB/s</td>
</tr>
<tr>
<td>ABL MLC PROGRAMMING FULL-SEQUENCE</td>
<td>34MB/s</td>
</tr>
<tr>
<td>16 Gb ABL DIE SIZE</td>
<td>182mm²</td>
</tr>
</tbody>
</table>
Summary, 3 bit per cell

<table>
<thead>
<tr>
<th>READ THROUGHPUT</th>
</tr>
</thead>
<tbody>
<tr>
<td>BY 8 I/O</td>
</tr>
<tr>
<td>ABL MLC PROGRAMMING</td>
</tr>
<tr>
<td>FULL-SEQUENCE</td>
</tr>
<tr>
<td>16 Gb D3 DIE SIZE</td>
</tr>
</tbody>
</table>
ABL is the architecture of the future

• The ABL architecture provides the fastest MLC program-throughput reported yet in 56nm and closes the gap between SLC and MLC

• By means of Full-Sequence programming the high voltage exposure is halved and the Bit Line interaction is minimized for maximum endurance

• The technology challenges of 43nm and beyond are well served by this architecture

• ABL is the architecture of choice for the next generations and 3 bit per cell encoding
Conclusions

• A new All Bit Line architecture boosts the MLC program-throughput of a 16Gb NAND by 240%

• With hierarchical column block architecture and a two-step redundancy access, a 50% faster internal I/O rate is achieved, even for lower power supply

• ABL provides an energy efficient system with an improved overall performance

• Lower die size when compared to a four plane “Conventional” chip of close performance
Acknowledgements

• The author wants to specially thank George Samachisa and Yan Li for their support.

• The circuits presented here were made possible by the work of many engineers from SanDisk and 1)Toshiba:

  Long Pham, Farookh Moogat, Siu Chan, Binh Le, Shouchang Tsao, Tai-Yuan Tseng, Khanh Nguyen, Jason Li, Jayson Hu, Cynthia Hsu, Fanglin Zhang, Teruhiko Kamei, Hiroaki Nasu, Phil Kliza, Khin Htoo, Jeffrey Lutze, Yingda Dong, Masaaki Higashitani, Junnhui Yang, Hung-Szu Lin, Vamshi Sakhamuri, Jong Hak Yuh, Alan Li, Feng Pan, Sridhar Yadala, Subodh Taigor, Kishan Pradhan, James Lan, James Chan, Takumi Abe1, Yasuyuki Fukuda1, Hideo Mukai1, Koichi Kawakami1, Connie Liang, Tommy Ip, Shu-Fen Chang, Jaggi Lakshmipathi, Sharon Huynh, Dimitris Pantelakis, Seungpil Lee, Yupin Fong, Tien-Chien Kuo, Jong Park, Tapan Samaddar, Hao Nguyen, Man Mui, Emilio Yero, Gyuwan Kwon, Jun Wan, Tetsuya Kaneko1, Hiroshi Maejima1, Hitoshi Shiga1, Makoto Hamada1, Norihiro Fujita1, Kazunori Kanebako1, Eugene Tam, Anne Koh, Iris Lu, Calvin Kuo, Trung Pham, Jonathan Huynh, Qui Nguyen, Hardwell Chibvongodze, Mitsuyuki Watanabe, Ken Oowada, Grishma Shah, Byungki Woo, Ray Gao, Patrick Hong, Liping Peng, Debi Das, Dhritiman Ghosh, Vivek Kalluru, Sanjay Kulkarni, Mehrdad Mofidi, Khandker Quader, Chi-Ming Wang