Solid-state drive (SSD) and memory system innovation

Dec. 11, 2008
Ken Takeuchi
Dept. of Electrical Engineering and Information Systems, University of Tokyo
E-mail: takeuchi@lsi.t.u-tokyo.ac.jp
http://www.lsi.t.u-tokyo.ac.jp

Outline

- SSD, Memory System Innovation
- NAND Overview
- NAND Circuit Design
- SSD Overview
- NAND Controller Design
- Operating System for SSD
- Green IT with SSD
- Summary

Definition of SSD

- SSD: Solid-State Drive
  - Mass storage to replace HDD of PC/Enterprise Server.
  - Small, robust, low-power and high performance.
- SSD consists of NAND Flash Memory and NAND controller(+RAM)

NAND Flash Memory and SSD Market

- PC expected as an emerging application
Memory System Bottleneck

CPU registers (<1ns)

SRAM (<1ns)

DRAM (10ns)

Big Gap

HDD (10ms)

SLC NAND as Cache of HDD

CPU registers (<1ns)

SRAM (<1ns)

DRAM (10ns)

SLC NAND (20us)

HDD (10ms)

Future Memory System

CPU registers (<1ns)

S/DRAM (<1ns)

DRAM (10ns)

NAND Controller

DRAM (10ns)

1bit/cell NAND (20us)

2-4bit/cell NAND (1~10ms)

SSD

Future Direction: Vertical Integration

History of NAND Flash Memory System

Application Software

File System (OS)

NAND Controller

Bad Block Management

Wear-leveling

ECC

NAND Flash Memory

Future Block Abstracted SSD

MP3 Player

SD Card

USB Memory

Go vertical integration to improve system-level performance.

Smart Media

NAND Flash Memory

Future Block

Abstracted SSD

Key Challenge of SSD

- Need to improve device reliability such as endurance, data retention, and disturb.

- Require co-design of NAND and NAND controller circuits to best optimize both NAND and NAND controllers.

- OS/Computer architecture innovation essential.

Outline

- SSD, Memory System Innovation
- NAND Overview
- NAND Circuit Design
- SSD Overview
- NAND Controller Design
- Operating System for SSD
- Green IT with SSD
- Summary
NAND Flash Memory

43nm 16Gb NAND

NAND flash memory chip
Memory circuit
Memory cell: Floating Gate-FET

Page & Block of NAND Flash Memory

Page: program/read unit
Block: Erase unit

Memory cells are sandwiched by select gates.
Contactless structure: ideal 4F<sup>2</sup> cell size

Top View of NAND Flash Cell Array

Simple structure: High scalability, High yield

MLC vs. SLC

SLC: Single-level cell or 1bit/cell
MLC: Multi-level cell or >2bit/cell
2bit/cell: Long production record since 2001
3bit/cell or 4bit/cell: Will be commercialized this year.
Most existing SSD uses SLC. MLC based SSD is commercialized this year.

NAND Operation Principle

Read

- Bit-line (0.8V-4.5V)
- Vread (4.5V)
- Selected word-line (Read voltage : 0V)
- Vread (4.5V)
- Vread (4.5V)

After precharging, bit-lines are discharged through the memory cell.
Unselected cells are biased to the pass voltage, Vread.
Small cell read current (~1uA) → Slow random access (~50us)
Serial access: 30-50ns → Fast read = 20-30MB/sec

Program: Electron injection

- Channel-FN tunneling
- High reliability
- Low current consumption (~pA/cell)
- Page based parallel program
  Typical page size: 2-4kB

Erase: Electron ejection

- 20V 20V
- 0V
- 0V 18V
- 0V

S. Aritome, IEDM 1990, pp.111-114.
**NAND Operation Principle (Cont')**

- **Page based parallel programming**
  - Memory cell array
  - All memory cells in a page are programmed at the same time.
  - Program speed = \( \frac{\text{Page size}}{\text{Programming time}} \)
    - \( 8\text{KByte} / 800\text{us} \)
    - \( 10\text{MByte/sec (56nm MLC)} \)

  T. Tanaka, Symp. on VLSI Circuits 1990, pp.105-106.


**Outline**

- SSD, Memory System Innovation
- NAND Overview
- NAND Circuit Design
- SSD Overview
- NAND Controller Design
- Operating System for SSD
- Green IT with SSD
- Summary

**NAND Circuit Design**

- **Random Access**
  - High Speed Programming
  - High Speed Read

- **Sequential Access**
  - High Speed Programming
  - High Speed Read

**Random Access : High Speed Programming**

- **Bit-by-bit Program Verify Scheme**

  During the verify-read, the program data in the page buffer is updated so that the program pulse is applied ONLY to insufficiently programmed cells.


  Achieve both fast programming and precise Vth control.

**Random Access : High Speed Programming (Cont')**

- **Incremental Program Voltage Scheme**

  - Word-line waveform
  - Program characteristics
  - Vth shift is constant at \( V_{PP} \)


**Random Access : High Speed Programming (Cont')**

- **Problems of MLC programming**

  - Two bits in a cell are assigned to two column addresses.
  - 3 operations ("1"-, "2"-, and "3"-program) required.
  - Long programming.
Random Access: High Speed Programming (Cont')

**Solution: Multi-page Cell Architecture**

- 2nd page program
- Number of memory cells
- Two bits in a cell are assigned to two row addresses.
- In average, 1.5 operations.
- Twice faster than conventional scheme.

Random Access: High Speed Programming (Cont')

**Program Voltage Optimization**

- WL0, 31: Higher capacitive coupling with word-lines. Initial program voltage is set lower.
- Optimized program voltage accelerates the programming.

Random Access: High Speed Programming (Cont')

**Problems: FG-FG interference**

- FG-FG coupling shifts the Vth of a memory cell as the neighboring cell are programmed.
- To tighten the Vth distribution, $\Delta V_{pgm}$ is decreased, causing a slow programming.
- The Vth modulation becomes significant as the memory cell is scaled down.

Random Access: High Speed Read (Cont')

**Solution: Multi-page Cell Architecture**

- 2nd page program
- Number of memory cells
- Two bits in a cell are assigned to two row addresses.
- In average, 1.5 operations.
- Twice faster than conventional scheme.

Random Access: High Speed Read

**Problems of MLC read**

- Two bits in a cell are assigned to two column addresses.
- 3 operations ("1"-, "2"- and "3"-read) required.
- Long random read.
Sequential Access : High Speed Programming

- Parallel Operation
  - Increase page size
  - Multi-page operation
  - Multi-chip operation (Interleaving)
    - To be discussed in “NAND Controller Circuit Design” section
- Pipeline Operation
  - Write/Read Cache
  - Cache Page Copy

Parallel Operation : Increase Page Size

- Page size trend
  - By increasing the word-line length, the page size has been extended to increase the write and read throughput.

- But, the large page size also causes problems.
  - Noise issue due to the large RC delay of a word-line

Parallel Operation : Increase Page Size (Cont’)

- Problems : SG-WL noise
  - [Conventional read/verify-read]

Parallel Operation : Increase Page Size (Cont’)

- Problems : WL-WL noise
  - [Conventional verify-read]

Parallel Operation : Increase Page Size (Cont’)

- Solution
  - Raise neighboring SG BEFORE bit-line discharge

Parallel Operation : Increase Page Size (Cont’)

- Solution
  - Unselected WLs are fixed at 5.5V
  - Inter-WL noise eliminated

Programming : 21% faster
Parallel Operation : Multi-page Operation

- Multi-page operation
  - Operate multi-page simultaneously to increase the write/read throughput.

Pipeline Operation : Write/Read Cache

- Pipelining of data-in/out & cell read/write
  - Implement data cache in NAND
  - Input/output data to the data cache during cell read/program

Outline

- SSD, Memory System Innovation
- NAND Overview
- NAND Circuit Design
- SSD Overview
- NAND Controller Design
- Operating System for SSD
- Green IT with SSD
- Summary

SSD Performance

- Random access
  - OS changes such as directory entry and file system metadata
  - Application S/W change
  - 50% of data is < 4KB.
  - Random access mainly decides the performance of PC.

- Sequential access
  - Boot
  - Hibernation

SSD Performance (Cont')

- Random access
  - Read : SSD with SLC and MLD has a great advantage over HDD.
  - Write : SSD still has a performance advantage. Write performance can be an issue in the future if the NAND performance degrades by scaling the memory cell or increasing the number of bits per cell.

SSD Performance (Cont')

- Sequential access

<table>
<thead>
<tr>
<th></th>
<th>NAND : Single chip operation</th>
<th>NAND : 4 chip interleaving</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Read</td>
<td>Write</td>
</tr>
<tr>
<td>NAND (SLC)</td>
<td>25us</td>
<td>300us</td>
</tr>
<tr>
<td>NAND (MLC)</td>
<td>50us</td>
<td>800us</td>
</tr>
<tr>
<td>HDD</td>
<td>3ms</td>
<td>3ms</td>
</tr>
</tbody>
</table>

Erase are hidden by operating the erase during the idle period.

- Read : SSD with SLC and MLD has a great advantage over HDD.
- Write : SSD still has a performance advantage. Write performance can be an issue in the future if the NAND performance degrades by scaling the memory cell or increasing the number of bits per cell.
SSD Performance (Cont')

- Slow random write problem
  - Page: program/read unit
  - Block: Erase unit

Bitline

- In case a part of the block is over-written, a block copy operation is performed.

Garbage Collection & Slow Random Write

- System performance degradation of a large block

\[ \text{System performance degradation} = \left( T_{\text{Cell read}} + T_{\text{Data out}} + T_{\text{ECC}} + T_{\text{Cell program}} \right) \times (\text{# of pages per block}) \]

\[ = 125\text{ms} \]

Pipeline Operation: Cache Page Copy

- Solution: Fast block copy

Step 1: Old block
Step 2: New block
Step 3: Old block
Step 4: New block

Solution for Slow Random Write

- Fast pipeline block copy operation
- Smaller block size (All bit-line architecture)
- Page based data allocation
  - Not to overwrite an old page but write data to an empty page.
  - Change the logical-physical address table.

SSD Power Consumption

- Power consumption
  - NAND: Single chip operation
  - NAND: 4 chip interleaving
    - Read
    - Write
    - NAD (SLC)
    - NAD (MLC)
    - HDD

- In SSD, additional current (~100mA) are consumed in the NAND controller, RAM and IO.
- Actual Power Consumption
  - HDD
  - SSD

- In all modes, the power consumption of SSD is smaller than HDD.
SSD Reliability

- SSD is robust.
- No mechanical parts.
- Need to be careful in PC/server application.
- Portable consumer electronics application (Digital still cameras, MP3 players, Camcorders).
- Effective data retention time << 10 years.
- Data quickly transferred to PC or DVD through USB drive and memory cards.
- Most probably data backup in PC.
- PC/Enterprise server application.
  - Higher reliability required w/o backup.
  - Need longer data retention time: 5-10 years.

SSD Reliability (Cont')

- Failure mechanism of NAND
  - Program disturb
    During programming, electrons are injected to unselected memory cells.
  - Read disturb
    During read, electrons are injected to unselected memory cells.
  - Write/Erase endurance & Data retention
    As the Write/Erase cycles increase, damage of the tunnel oxide causes a leakage of stored charge.

SSD Reliability (Cont')

- "Classic" program disturb

Program inhibit
  Bitline (Vcc)
  Vpgm (18V)
  Vpass (10V)

Vpgm disturb cell
  18V
  ~8V

Vpass disturb cell
  10V
  0V

Both selected and unselected cells suffer from the disturb.

SSD Reliability (Cont')

- "Modern" program disturb

Hot carriers generated at the select gate edge inject into the memory cell causing a Vth shift.
- The Vth shift can be reduced by increasing SG-WL space.

SSD Reliability (Cont')

- "Modern" program disturb (Cont')

Select Tr. Dummy Tr. WL0

The Vth shift can be reduced by adding dummy WL.

SSD Reliability (Cont')

- Read disturb

Weak program bias condition
- Unselected word-lines suffer from the read disturb.
SSD Reliability (Cont')

- Program disturb and read disturb summary
  - Program disturb and read disturb is a “bit error” not a “burst error”.
  - Two bits in MLC are assigned to different pages.
  - Even if one MLC cell fails, one bit in two pages fails.
  - ECC (Error correcting code) effectively corrects the bit error.
  - Existing ECC corrects 4-12 bit errors per 512Byte sector.

- Page assignment of MLC
  - 2-level cell
  - 4-level cell

- Program disturb and read disturb summary
  - ECC (Error correcting code) effectively corrects the bit error.

SSD Reliability (Cont')

- Write/Erase Endurance & Data Retention
  - Endurance: how many times data are written
  - Data retention: how long the data remains valid
  - Clear correlation between endurance and data retention
  - Damages to the tunnel oxide during write and erase cause the data retention problems.
  - Traps are generated during write and erase.
  - The unlucky cell with traps results in a leakage path, causing the charge transfer.
  - The leakage current is called SILC (Stress Induced Leakage Current).
  - To guarantee data retention, Write/Erase cycles are limited to 100K (SLC) or 10K (MLC).

SSD Reliability (Cont')

- 100K (SLC) or 10K (MLC) W/E cycles acceptable?
  - W/E cycles estimation for PC
    - 32GB SSD
    - Usage scenario: 2-5GB/day (1)
    - Service for 5 years
    - 100% efficient wear leveling
    - (365 days/year) x 5 years / (32GB / 2-5GB/day)
    - = 114-285 W/E cycles
    - 114-285 cycles are far below the NAND limitation of 100K for SLC or 10K for MLC.
    - Actual W/E cycles are higher for the file management such as garbage collection.

Outline

- SSD, Memory System Innovation
- NAND Overview
- NAND Circuit Design
- SSD Overview
- NAND Controller Design
- Operating System for SSD
- Green IT with SSD
- Summary

SSD SW Architecture

- OB
  - File system
  - Low level driver
  - ATA I/F

- SSD
  - NAND I/F
  - Flash Translation Layer (FTL)
    - Bad block management
    - Wear-leveling
    - Address translation from logical address to physical address of NAND
    - Interleaving
    - ECC

- NAND I/F
  - NAND Flash Memory

HW Architecture

- Block diagram (Single channel)
  - HDD-like architecture: DRAM buffer to hide NAND random access
  - High power consumption
  - High cost

Outline

- SSD, Memory System Innovation
- NAND Overview
- NAND Circuit Design
- SSD Overview
- NAND Controller Design
- Operating System for SSD
- Green IT with SSD
- Summary

Outline

- SSD, Memory System Innovation
- NAND Overview
- NAND Circuit Design
- SSD Overview
- NAND Controller Design
- Operating System for SSD
- Green IT with SSD
- Summary

Outline

- SSD, Memory System Innovation
- NAND Overview
- NAND Circuit Design
- SSD Overview
- NAND Controller Design
- Operating System for SSD
- Green IT with SSD
- Summary
HW Architecture (Cont')

- Block diagram (Multi-channel)

DRAM eliminated:
- Random access of NAND is faster than HDD.
- Low power consumption
- Low cost
- Multi-channel
- Parallel operation
- High bandwidth

High Speed Technology

- Interleaving: Sequential Parallel Write


- Data loading time
- One page program time
- Automatic write interleaving engine without SW intervention

- 2-channel 4-way interleaving
- Max write throughput: 80MB/sec for MLC.
- HW driven automatic operation.

High Reliability Technology

- Wear-leveling
  - Problem
    - Write/Erase cycle of NAND is limited to 100K for SLC and 10K for MLC.
  - Solution
    - Write data to be evenly distributed over the entire storage.
    - Count # of Write/Erase cycles of each NAND block.
    - Based on the Write/Erase count, NAND controller re-map the logical address to the different physical address.
    - Wear-leveling is done by the NAND controller (FTL), not by the host system.

Example of wear-leveling

- If the block is occupied with old data, data is programmed to a new block.
- If there is no free block, the invalid block are erased.

Block 1
Block 2
Block 3
Block 4
Invalid

Rewrite old file to an empty block

Write new file to an empty block

Red: Static data such as system data
Blue: Dynamic data such as user data

Physical block address

High Reliability Technology (Cont')

- Static data
  - Data that does not change such as system data (OS, application SW).
- Dynamic data
  - Data that are rewritten often such as user data.

Dynamic wear-leveling
- Wear-level only over empty and dynamic data.

Static wear-leveling
- Wear-level over all data including static data.
High Reliability Technology (Cont’)

- **Static wear-leveling**
  - **Write/Erase count**
  - Red: Static data such as system data.
  - Blue: Dynamic data such as user data
  - Wear-level more effectively than dynamic wear-leveling.
  - Search for the least used physical block and write the data to the location. If that location is empty, the write occurs normally.
  - Contains static data, the static data moves to a heavily used block and then the new data is written.

High Reliability Technology (Cont’)

- **Bad Block Management**
  - Program/Erase characteristics vs. endurance
  - As the Write/Erase cycles increase, erase failure occurs, resulting in a bad block.
  - The NAND controller detects and isolates the bad block.

High Reliability Technology (Cont’)

- **ECC (Error Correcting Code)**
  - To overcome read disturb, program disturb and data retention failure, ECC have to be applied.
  - Since failure pattern is random, BCH is sufficient.
  - Existing NAND controller can correct 4-12bit error per 512Byte sector.
  - NAND with embedded ECC is also published.

Outline

- SSD, Memory System Innovation
- NAND Overview
- NAND Circuit Design
- SSD Overview
- NAND Controller Design
- Operating System for SSD
- Green IT with SSD
- Summary

Why OS?

- **Motivation**
  - Existing OS is optimized for magnetic drives.
  - Current SSD based PC uses the conventional OS and just replace HDD with SSD.
  - To achieve the best performance and reliability of SSD, OS especially file system should be optimized.
  - Windows 7 will treat SSD differently from HDD.

New Memory System: NAND/HDD Combo

- **NAND as a cache**
  - Intel Robson
  - Microsoft Ready Boost
- **Multi-drive of NAND/HDD**
  - SanDisk Vaulter Disk
  - NAND: OS data
  - HDD: User data
  - Temporary solution until NAND cost becomes comparable with HDD cost.
MLC/SLC Hybrid SSD

Future Direction: Hybrid SSD with SLC and MLC
- Concept: Right device for the right use.
- Enjoy the Benefit of both SLC and MLC.
- SLC: Fast and highly reliable but low capacity.
- Use SLC as a cache or system data storage.
- MLC: Large capacity but slow. Use MLC as user data storage.
- OS support essential: SSD does not know the contents of the file.

Performance Optimization

Sector size optimization
- Minimum write/read unit of NAND is a page.
- Typical page size is 4-8KByte.
- A page is written only ONCE to avoid the program disturbance.
- With current OS having 512Byte sector, one sector write wastes >80% of data in a page.

Page Size Trend
- As the page size increases as NAND is shrinking, larger sector size such as 64KByte or 128KByte is required.

Reliability Optimization

Enhanced Write Filter (Windows Embedded)
- Decrease write/erase cycles of NAND, extending the NAND lifetime.
- Control the file allocation to store frequently rewritten file in DRAM and not to access NAND.
- Enhanced Write Filter (EWF) is located between file system and low level driver interfacing with SSD.
- OS/Application SW support essential. Again, SSD does not know the contents of the file.

Reliability Optimization (Cont’)

SMART
(Self-Monitoring, Analysis and Reporting Technology)
- Monitor the storage and report/predict the failure.
- SMART for HDD is NOT smart because it is very difficult to predict the mechanical failure.
- SMART for SSD can be really smart.
- Product lifetime can be predicted because the failure rate is highly correlated with the write/erase cycles.
- Predict the SSD lifetime by monitoring the write/erase cycles and replace SSD before the fatal failure occurs.
Green IT : Power Crisis of Data Center

- Data through internet is increasing drastically.
- In the U.S, power consumption at the data center doubled during last 5 years. (5 nuclear power plants!)
- In 2025, the data increases by 200 times and the power consumption increases by 12 times.

Problems of NAND Flash Memory

- Reliability
  - Low write/erase cycles: Currently <10K cycles (MLC) and decreasing as scaling down memory cells.
  - >100K cycles required

- Power Consumption
  - Because of the scaling, the parasitic capacitance increases and the power consumption doubles.
  - Low power memory device required

- Capacity
  - Currently Gbyte → TByte required

Operation Current Trend of NAND

- In the scaled VLSIs, most power is consumed to charge and discharge signal-lines.
- Inter signal-line capacitance, $C_{\text{wire-wire}}$ drastically increases to keep the low signal-line resistance.

Fe(Ferroelectric)-NAND Flash Memory

- NAND Flash Memory w. Ferroelectric Transistor
  - Scalable below 20nm
  - Low voltage/power operation: 20V→5V
  - Write/Erase cycles: 10K cycles→100M cycles
  - Most suitable for data server application
Operation Principle of Fe-NAND Flash

- Low voltage operation

10 Year Data Retention

- 10 years data retention

Excellent W/E Cycles up to 100M

- Fe-NAND

Co-design of NAND and Controller Circuits

- By co-designing both NAND and NAND controller circuits, the power consumption of SSD is reduced by 60%.

Co-design of NAND and Controller Circuits

- Low Power Circuit Technology
  - Selective bit-line precharge scheme
  - Advanced source-line program

- Low Noise Circuit Technology
  - Intelligent interleaving
SSD Write Performance

- Interleaving: write $N$ NAND chips in parallel
  $$\text{Performance}_{\text{SSD}} = N \times \text{Performance}_{\text{NAND}}$$
- Limitation of $N$
  $$N \times I_{\text{NAND}} < I_{\text{cc}_\text{constraint}}$$
- Key NAND design issue:
  Decrease $I_{\text{NAND}}$ to maximize $N$

NAND Bit-line Capacitance Trend

- Inter bit-line capacitance, $C_{\text{BL-BL}}$, drastically increases to keep the low bit-line resistance.

Current & Performance Trend

- NAND operation current
- SSD performance

C_{\text{bit-line}} \uparrow \Rightarrow I_{\text{NAND}} \uparrow \Rightarrow N \downarrow \Rightarrow \text{Performance}_{\text{SSD}} \downarrow

- Low power circuit of NAND required.

Multi-level Cell Program

- Algorithm
  1. Data load
  2. Program pulse
  3. Verify read
  4. All cells programmed?
  5. End

Conventional Verify Read

- All bit-lines are precharged irrespective of the program data.
- Page based program
  - 8KByte (Page size) bit-lines are precharged.
  - Total bit-line capacitance > 200nF!

Selective Bit-line Precharge Scheme

- During verify, precharge bit-line based on the program data in the page buffer.
- Skip unnecessary bit-line precharge and save current during verify.
- Area overhead < 1%
1st Page Program ("A"-verify)

Number of memory cells
Erased state
"A"  "B"  "C"

<table>
<thead>
<tr>
<th></th>
<th>[Conventional]</th>
<th>[Proposed]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st page program</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2nd page program</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bit-line</th>
<th>precharge</th>
<th>discharge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vbit-line</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Vth

"A"-program complete / incomplete
Program inhibit

"A"-program in-complete

Program inhibit

[Source-line program]

<table>
<thead>
<tr>
<th>Bit-line (2.5V)</th>
<th>Source-line (1V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SGD (4.5V)</td>
<td>Inhibit voltage</td>
</tr>
<tr>
<td>Virgin (10V)</td>
<td></td>
</tr>
<tr>
<td>Vpass (10V)</td>
<td></td>
</tr>
</tbody>
</table>

Source-line (1V) 2S G, 6 4C G

Cbit-line = Cwire(bit-line) + Cjunction
Csource-line = Cwire(source-line) + Cjunction

Capacitance Comparison

Demonstrated in 0.25um NAND

Cbit-line >> Cjunction
Csource-line >> Cwire(source-line)

Source-line Program (VLSI'99)

- Save current during program pulse.
- Bias 2.5V from a low capacitance source-line.
- Low voltage swing for a high capacitance bit-line

Result – Selective BL precharge

- 23% current reduction.
- 50% performance improvement.

SSD Performance [MByte/sec]

Feature size [nm]

Selective BL precharge

Source-line (2.5V)

Vpass (10V)

Source-line (1V)

Demonstrated in 0.25um NAND

### Advanced Source-line Program

- Total source-line capacitance in a chip exceeds 20nF for sub-50nm NAND.
- Hierarchical source-line structure
- Only metal layout change; No area overhead

### Read Operation

- All Source-line switch: ON
- Minimize the source-line resistance.
- Suppress the source-line noise.

### Program Operation

- Only one of 16 sub-arrays activated
- Source-line capacitance: 90% reduced

### Result – Advanced SL program

- 60% current reduction.
- 250% performance improvement.

### Intelligent Interleaving

- Disperse the current peak and avoid the power supply noise.
- Bit-line precharge & charge pump ramp-up cause >100mA current peak.

### Control Circuit

- Introduce Power Detect (PD) signal.
- When a NAND starts bit-line precharge & charge pump ramp-up, PD becomes low.
- NAND controller issues a write command when PD is high and NAND is ready.
Summary: Co-design of NAND and Controller Circuits

- Co-design of NAND and NAND controller.
  - To improve SSD speed, decrease the NAND current and maximize # of NAND operated in parallel.
- Two low power circuit technologies proposed.
  - Selective bit-line precharge scheme
  - Advanced source-line program
  - 60% current reduction.
  - 250% SSD performance improvement.
- Intelligent interleaving realizes highly reliable and high-speed SSD.

3D-integrated SSD

- First demonstration of 3D-integrated SSD.
- With smart Mix & Match, the power decreases by 70%.

Summary

- New Memory System
  - SLC/MLC Hybrid SSD solves the system bottleneck.
- Emerging Market: Power Crisis at data center
  - SSD is expected to save power at data center.
- Device, circuit and OS innovation required.
  - Co-design of NAND and NAND controller circuits
  - OS optimization such as sector size optimization
  - Fe(Ferroelectric)-NAND flash memory device
  - 3D-integrated SSD circuits

Thank you!

E-mail : takeuchi@lsi.t.u-tokyo.ac.jp
http://www.lsi.t.u-tokyo.ac.jp