SSD Characterization:
From Energy Consumption's Perspective

Balgeun Yoo, Youjip Won
Seokhei Cho, Sooyong Kang

Jongmoo Choi

Sungroh Yoon

Hanyang University
Dankook University
Korea University
Outline

- Motivation
- SSD Organization and Energy Consumption
- Channels, Ways and Clusters
- Case Study
- Power Budget
- Conclusion
Motivation

Understanding of Internal Mechanism of Storage Device is very important!

- Hard disk
  - Sector layout: cylinder serpentine vs. surface serpentine vs. hybrid serpentine
  - Number of zones
  - Degree of track skew
  - Disk scheduling algorithm
Characterizing HDD: Sector Layout

- Jongmin Gim et al, ACM ToS 6, 2 (July 2010)

**Diagram:**
- Traditional Spindle
- Surface serpentine
- Cylinder serpentine
- Zone rewind

**Graph:**
- Hitachi HDT7725032VLA360 for 2000 tracks
- X-axis: track number
- Y-axis: time (ms)
HDD Characterization is via measuring **Seek time** and **Rotational Latency**.

**Characterizing SSD... what do we use?...**
How do we figure out the internals of SSD?

- What is available
  - the number of channels
  - the number of chips/packages per channels

- What is not available?
  - Sector placement, Garbage collection algorithm
Operations on NAND Flash Cell

Program:

- Voltage range: 18~35V
- Control Gate
- Floating Gate
- Oxide Layer
- Source
- Drain
- Substrate

Erase:

- Control Gate
- Floating Gate
- Oxide Layer
- Source
- Drain
- Substrate
Operations on NAND Flash Cell
SSD Characterization

We will use “Energy Consumption”
SSD Organization and Energy Consumption
SSD Organization

Host

I/O Physical Interface

Host Interface

Bus

DRAM

Controller

NAND Flash Memory

Controller

DRAM

Bus

Host

I/O Physical Interface

Host Interface

NAND Flash Memory
Characteristics of SSD Behavior

- Page Write
- Write Complete

Diagram of SSD behavior with yellow and light blue boxes.
Channels, Ways, and Clusters
Writing 4 pages: 1 Channel X 1 Way

- Page Write
- Write Complete

![Diagram showing the process of writing 4 pages with channel X 1 way, with yellow blocks representing page write and blue blocks representing write complete.](image_url)
Way switch vs. Channel switch Delay

- **Page Write**
- **Write Complete**

### Way Switch
- 

### Channel Switch
- 

**Current (mA)**

**P**

**D**

**C**

**time**
Writing 4 pages: 1 Channel X 2 Way

- Page Write
- Write Complete

Diagram showing the process of writing 4 pages through 1 channel with 2 ways, indicating the current flow (mA) over time.
Writing 4 pages: 2 Channel X 1 Way

- =Page Write
- =Write Complete

Current (mA) vs. time

- C
- D
- P

Youjip Won
Hanyang University
Writing 4 pages: 2 Channel X 2 Way

- Yellow: Page Write
- Blue: Write Complete

Diagram shows a timing chart with current (mA) and time axes, illustrating the process of writing 4 pages through 2 channels with 2 ways.
Cluster: Write Unit of SSD

- Page Write

4KB Write

8KB Write

= Page Write

...
Measurement

- Oscilloscope (TDS3032)
- High resolution current probe (TCP202)
- Current probe to power line ($V_{dd}$) of the SSD
- Sampling interval (10 samples MA): 10 $\mu$s

<table>
<thead>
<tr>
<th>Model</th>
<th>Vendor</th>
<th>Size</th>
<th>Channels</th>
<th>DRAM Size</th>
<th>Package</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>X25M</td>
<td>Intel</td>
<td>80GB</td>
<td>10</td>
<td>16MB</td>
<td>20</td>
<td>MLC</td>
</tr>
<tr>
<td>MXP</td>
<td>SAMSUNG</td>
<td>128GB</td>
<td>8</td>
<td>128MB</td>
<td>16</td>
<td>MLC</td>
</tr>
</tbody>
</table>
## Case Study: Intel X25M

<table>
<thead>
<tr>
<th>Specification</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capacity</td>
<td>80GB</td>
</tr>
<tr>
<td>No. of Channels</td>
<td>10</td>
</tr>
<tr>
<td>Packages/Channel</td>
<td>2</td>
</tr>
<tr>
<td>Package</td>
<td>4 GB</td>
</tr>
</tbody>
</table>

![Diagram of Intel X25M NAND packages](image-url)
Case Study: Intel X25M

- IO Size: 4KB to 80KB

Channel switch = 30 μsec
NAND programming: 17 mA
Simple Round Robin

- 600 μsec
- Increase 20 steps
- 17 mA
- 100 mA
Case Study: Intel X25M

- 4KB to 80KB Sequential Write

= Page Write

Channel Switch = 30 μsec
Case Study: Intel X25M

80KB to 160KB Write

mA

Time (ms)

80 KByte

84 KByte

HotStorage ’11, Portland, OR
Case Study: Intel X25M

=1 Page Write
=Write Complete

MUX

=Write Complete

Flow

Current

Youjip Won
Hanyang University

HotStorage '11, Portland, OR
### Case Study: SAMSUNG MXP

**Capacity**: 128GB  
**No. of Channels**: 8  
**Packages/Channel**: 2  
**Package**: 8 GB

---

![Diagram showing NAND 8GB packaging structure]
Case Study: SAMSUNG MXP

The graph shows the peak current for different byte sizes over time. The peak current values are 50mA and 100mA.
Case Study : SAMSUNG MXP

Same Peak Current

mA

Time(ms)

36Kbyte
40Kbyte
44Kbyte
48Kbyte
52Kbyte
56Kbyte
60Kbyte
64Kbyte
Case Study: SAMSUNG MXP

Same Peak Current

mA

Time (ms)

68Kbyte
72Kbyte
86Kbyte
80Kbyte
84Kbyte
88Kbyte
92Kbyte
96Kbyte
Case Study: SAMSUNG MXP

Cluster Size of MXP is 32KB

Same Peak Current

Cluster Size of MXP is 32KB
What we still do not know of

Peak of 128 Kbyte < Peak Current of 124 KByte
Forth Coming Problem in Multi-channel SSD

<table>
<thead>
<tr>
<th>Model Name</th>
<th>Release</th>
</tr>
</thead>
<tbody>
<tr>
<td>SATA3035 (Mtron)</td>
<td>2008.01</td>
</tr>
<tr>
<td>Vertex (OCZ)</td>
<td>2009.03</td>
</tr>
<tr>
<td>Vertex2 (OCZ)</td>
<td>2010.07</td>
</tr>
<tr>
<td>REVO Drive X2 (OCZ)</td>
<td>2011.01</td>
</tr>
</tbody>
</table>
Forth Coming Problem in Multi-channel SSD

- 10 channel: peak 500 mA
- 16 Channel: peak 800 - 900 mA → SSD is no long Green.
- Further, excessive peak current can cause...
  - supply voltage drop
  - ground bounce
  - signal noise
  - black-out
  - Etc...
Power Budget

- 8 Channels X 1 Way
- 4 Channels X 2 Ways
- 2 Channels X 4 Ways
- 1 Channel X 8 Ways

Feasible Solutions

Feasible Region

Power

I/O Latency
## Summary

<table>
<thead>
<tr>
<th></th>
<th>X25M</th>
<th>MXP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cluster Size:</strong></td>
<td>4KB</td>
<td>32KB</td>
</tr>
<tr>
<td><strong>Programming:</strong></td>
<td>17 mA</td>
<td>35mA</td>
</tr>
<tr>
<td><strong>Peak Current:</strong></td>
<td>500mA</td>
<td>350mA</td>
</tr>
<tr>
<td><strong>Channel switch:</strong></td>
<td>30 μsec</td>
<td></td>
</tr>
<tr>
<td><strong>Energy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Performance</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Small Write</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Large write</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Standby current</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Summary

- Energy Consumption is very good tool to characterize SSD.
- For larger number of channels, peak current will soon be a significant issue.
- We introduce the notion of Power Budge to resolve this issue.
Related Works

- Dongkun Shin et al, NVRAMOS 2010 Spring
  
  Applied different workloads (random, sequential, etc) to SSDs and Measured the power consumption

- Laura M. Grupp et al, MICRO 2009

  Custom Board + Flash Memory

  Read

  Program

  Erase

  Basic Operation

  Power Consumption of Flash Memory Basic Operations
Related Works

- Euiseong Seo et al, HotPower'08
  
  Applied different workloads (random, sequential, etc) to SSDs with different request sizes (varied the file systems)
  Measured the Power Consumption

- Vidyabhushan Mohan et al, Date ’10

CACTI 5.3

developed a detailed power model for the NAND flash chip itself with CACTI 5.3
## Operations on NAND Flash Cell

<table>
<thead>
<tr>
<th>Model</th>
<th>Read (mA)</th>
<th>Write (mA)</th>
<th>Erase (mA)</th>
<th>Voltage (V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MD332B (Intel)</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>3.3</td>
</tr>
<tr>
<td>K9XXG08XXM (Samsung)</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>3.3</td>
</tr>
<tr>
<td>MT29FXXG08CXXBB (Micron)</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>3.3</td>
</tr>
</tbody>
</table>
Power Consumption of Storage Devices: System Boot

SSDs have various stand-by currents

Active current of HDD

stand-by current s

0 1 2 3 4 5 6 7
0 100 200 300 400 500 600 700 800 900
mA

Time(s)

Fujitsu 2.5” HDD
SAMSUNG MXP
OCZ Vertex
Hanamicon Forte Plus
Intel X-25M
FTL

LPN: Logical Page Number

PPN: Physical Page Number

PBN: Physical Block Number

Write to LPN = 3
Write to LPN = 5

Write to LPN = 3
Write to LPN = 5

Write to LBN = 1
Offset = 0

Write to LBN = 1
Offset = 2

FTL

Page FTL

Block FTL

Hybrid FTL
Trivia in Measurement Methodology
Trivia of Measurement Methodology

- Sampling interval should be small enough. (Read/Program Latency)
- Smoothing the data to filter out measurement noise.

![Diagram showing smoothing effect on data]

- Original data with high noise and large peak values.
- Smoothing reduces noise and decreases peak values.

**Example:**
- Original mA values: 400, 500, 600, 700, 800, 900, 1000
- Smoothed mA values: 150, 200, 250, 300, 350, 400, 450

**Time (ms):** 0, 2, 4, 6, 8, 10, 12

**mA:** 0, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500