Methodologies for Calculating SSD Useable Life

Gary Drossel
Western Digital
Takeaways

- Endurance and performance play against each other
- Usage model has the biggest impact on useful life
- NAND media becoming more differentiated
- SSD architectures vary greatly
- Endurance (longevity) and performance vs. cost
Factors Affecting SSD Service Life

SSD Technology

- Wear-Leveling
- ECC
- Write Amplification
- NVM Config
- NVM Specs
- Over Provisioning
- Performance

Usage Model

- Storage Budget
- Capacity
- Data Size
- Randomness of data
- Write Duty Cycle

Life (Years)
SSD Media Quadrilemma

More capacity and lower $/Gb → smaller cells, more bits/cell → less electrons/cell → lower reliability

Endurance
SNR Gain
Capacity
Reliability

SOC Architecture

Performance
“Fast Read, Slow Write, Slower Erase”
“Deliver Performance Via Parallelism”

Retention
“Keep Bits in Media”

NAND Management

Endurance
“Get Bits onto Media”
“Get Bits Back from Media”

Density (bits/cell)

“More Bits Deliver More Capacity & Lower Cost”
“More Bits Deliver Less Performance & Reliability”

Soft-Decision Multi-Level Decoding
Retention competes with endurance - both determine the reliability of the NAND

Different retention vs P/E cycles behavior Vendor A versus B
Data Retention

- JEDEC JESD47 Data Retention Test Specifications
  - 10 years for 10% cycled blocks
  - 1 year for 100% cycled blocks
- Stronger ECC allows longer data retention
  - Requires more ECC parity bytes, may not support all sector sizes
- NAND blocks need to constrain number of PE cycles
- Some discussions is exploring a 6 month data retention with 2x PE cycles spec for enterprise applications
Stronger ECC Improves Retention
Error Correction
NAND Future

- Variances in performance and reliability
  - Node-to-node, vendor-to-vendor
- Need SoC to shield host from these variances
  - Tight coupling and lots of test data

Source: WD Internal Research
**Fundamental SSD Variations**

- **Embedded** – 100’s of write IOPS
  - Minimize RAM (KBs) and power consumption with “decent” sequential r/w speeds
- **Add DRAM** – 1000’s of write IOPS
  - Data cache or manage mapping tables at the page level
    - Reduces WA significantly, increasing useful life
    - IOPS benefits when combined with overprovisioning
- **Trade-offs**
  - Increase power consumption and SSD complexity (data path protection, hold-up circuitry, etc.)
Over-Provisioning

- More capacity than is listed on the datasheet
- Can be used as simple “extra” capacity
  - 25% more capacity = 25% more life
- Can be used strategically to increase the service life even more
  - Strategic re-mapping to decrease write amplification
  - 25% over-provision decreases write amplification from 4 to 2 → 2x life for 25% more cost
Write Amplification

- The proportion of host writes to NAND writes
- Results from a mismatch of page and block sizes
- Definitions:
  - Page: Write size
  - Block: Erase size
  - 64 pages per block in NAND
- Varies dramatically based on SSD architecture
Write Amplification (Embedded)

- Ground rules for a write with no erase required
  - Can write to free space above
  - Can’t overwrite
  - Can’t write below

![Diagram showing page block and write amplification rules]

Storage Developer Conference 2009
© 2009 Insert Copyright information here. All rights reserved.
Write Amplification (Embedded)

1. Host writes to H
Write Amplification (Embedded)

1. Host writes to H (1:1)
Write Amplification (Embedded)

1. Host writes to H (1:1)
2. Host writes to A
1. Host writes to H (1:1)
2. Host writes to A
Write Amplification (Embedded)

1. Host writes to H (1:1)
2. Host writes to A (1:64)
Add DRAM - WA Management

Consider a snapshot at time $t$
Columns are blocks, letters are block names or numbers
Rows represent pages in a block
25% Over provision

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
<th>h</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>8</td>
<td>16</td>
<td>24</td>
<td>32</td>
<td>40</td>
<td>48</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>9</td>
<td>17</td>
<td>25</td>
<td>33</td>
<td>41</td>
<td>49</td>
<td>57</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>18</td>
<td>26</td>
<td>34</td>
<td>42</td>
<td>50</td>
<td>58</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>11</td>
<td>19</td>
<td>27</td>
<td>35</td>
<td>43</td>
<td>51</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>20</td>
<td>28</td>
<td>36</td>
<td>44</td>
<td>52</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>21</td>
<td>29</td>
<td>37</td>
<td>45</td>
<td>53</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>14</td>
<td>22</td>
<td>30</td>
<td>38</td>
<td>46</td>
<td>54</td>
<td>62</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>15</td>
<td>23</td>
<td>31</td>
<td>39</td>
<td>47</td>
<td>55</td>
<td>63</td>
<td></td>
</tr>
</tbody>
</table>

Blocks storing user data

Over Provision Blocks
Add DRAM - WA Management

<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
<th>h</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>8</td>
<td>16</td>
<td>24</td>
<td>32</td>
<td>40</td>
<td>48</td>
<td>56</td>
</tr>
<tr>
<td>1</td>
<td>9</td>
<td>17</td>
<td>25</td>
<td>33</td>
<td>41</td>
<td>49</td>
<td>57</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>18</td>
<td>26</td>
<td>34</td>
<td>42</td>
<td>50</td>
<td>58</td>
</tr>
<tr>
<td>3</td>
<td>11</td>
<td>19</td>
<td>27</td>
<td>35</td>
<td>43</td>
<td>51</td>
<td>59</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>20</td>
<td>28</td>
<td>36</td>
<td>44</td>
<td>52</td>
<td>60</td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>21</td>
<td>29</td>
<td>37</td>
<td>45</td>
<td>53</td>
<td>61</td>
</tr>
<tr>
<td>6</td>
<td>14</td>
<td>22</td>
<td>30</td>
<td>38</td>
<td>46</td>
<td>54</td>
<td>62</td>
</tr>
<tr>
<td>7</td>
<td>15</td>
<td>23</td>
<td>31</td>
<td>39</td>
<td>47</td>
<td>55</td>
<td>63</td>
</tr>
</tbody>
</table>

Blocks storing user data

After a write sequence of 3, 12, 5, 6 36, 37, 7, 63

Over Provision Blocks

<table>
<thead>
<tr>
<th>i</th>
<th>j</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
<tr>
<td>63</td>
<td></td>
</tr>
</tbody>
</table>
### Add DRAM - WA Management

#### Blocks storing user data

<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
<th>h</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>8</td>
<td>16</td>
<td>24</td>
<td>32</td>
<td>40</td>
<td>48</td>
<td>56</td>
</tr>
<tr>
<td>1</td>
<td>9</td>
<td>17</td>
<td>25</td>
<td>33</td>
<td>41</td>
<td>49</td>
<td>57</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
<td>18</td>
<td>26</td>
<td>34</td>
<td>42</td>
<td>50</td>
<td>58</td>
</tr>
<tr>
<td>3</td>
<td>11</td>
<td>19</td>
<td>27</td>
<td>35</td>
<td>43</td>
<td>51</td>
<td>59</td>
</tr>
<tr>
<td>4</td>
<td><strong>12</strong></td>
<td>20</td>
<td>28</td>
<td><strong>36</strong></td>
<td>44</td>
<td>52</td>
<td>60</td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>21</td>
<td>29</td>
<td><strong>37</strong></td>
<td>45</td>
<td>53</td>
<td>61</td>
</tr>
<tr>
<td>6</td>
<td>14</td>
<td>22</td>
<td>30</td>
<td>38</td>
<td><strong>46</strong></td>
<td>54</td>
<td>62</td>
</tr>
<tr>
<td>7</td>
<td>15</td>
<td>23</td>
<td>31</td>
<td>39</td>
<td>47</td>
<td><strong>55</strong></td>
<td><strong>63</strong></td>
</tr>
</tbody>
</table>

#### Over Provision Blocks

<table>
<thead>
<tr>
<th>i</th>
<th>j</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>36</td>
<td></td>
</tr>
<tr>
<td>37</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
</tr>
<tr>
<td>63</td>
<td></td>
</tr>
</tbody>
</table>

Garbage collection occurs
Add DRAM - WA Management

<table>
<thead>
<tr>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
<th>h</th>
<th>i</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>16</td>
<td>24</td>
<td>32</td>
<td>40</td>
<td>48</td>
<td>56</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>17</td>
<td>25</td>
<td>33</td>
<td>41</td>
<td>49</td>
<td>57</td>
<td>12</td>
</tr>
<tr>
<td>10</td>
<td>18</td>
<td>26</td>
<td>34</td>
<td>42</td>
<td>50</td>
<td>58</td>
<td>5</td>
</tr>
<tr>
<td>11</td>
<td>19</td>
<td>27</td>
<td>35</td>
<td>43</td>
<td>51</td>
<td>59</td>
<td>6</td>
</tr>
<tr>
<td>12</td>
<td>20</td>
<td>28</td>
<td>36</td>
<td>44</td>
<td>52</td>
<td>60</td>
<td>36</td>
</tr>
<tr>
<td>13</td>
<td>21</td>
<td>29</td>
<td>37</td>
<td>45</td>
<td>53</td>
<td>61</td>
<td>37</td>
</tr>
<tr>
<td>14</td>
<td>22</td>
<td>30</td>
<td>38</td>
<td>46</td>
<td>54</td>
<td>62</td>
<td>7</td>
</tr>
<tr>
<td>15</td>
<td>23</td>
<td>31</td>
<td>39</td>
<td>47</td>
<td>55</td>
<td>63</td>
<td>63</td>
</tr>
</tbody>
</table>

Blocks storing user data

Over Provision Blocks

Block a is erased and wear leveled

Garbage collection dynamically wear levels between blocks that contain frequently updated data and the over provision blocks
SSD Endurance Specifications

- Block level (cycles)
- Longterm Data Endurance (LDE)
- LifeEst corner case

“Your mileage may vary”
NAND Component Endurance

Data Retention (Years)

Endurance (Cycles)

- 6xnm SLC
- 5xnm SLC
- 3xnm SLC
- 5xnm MLC
Longterm Data Endurance (LDE)

-Specifies number of terabytes written
-Calculates time using GB per day

Example
-SSD rated at 40TBW
-Writing 40GB/day = 1,000 days = ~ 3 years
-Writing 20GB/day = 2,000 days = 6-7 years

BAPCo use case well documented for PC users
LifeEst Corner Case

- Based on parameters the SSD manufacturer controls
- Assumes **worst case** usage model with 100% write duty cycle
- Assumes worst case write amplification but can improve with hardware and firmware algorithms
LifeEst Corner Case

LifeEst
(SSD Technology) = SSD Endurance

SSD Endurance:
- ECC
- NVM Config
- NVM (20x)
- Wear-Leveling
- Over Provision

Unit Conversion Constant

Performance (100x) × Write Amplification (64x)

Life (Years) = LifeEst

LifeEst:
- Capacity (32x)
- Write Duty Cycle (100x)
Life Calculation Example

- **Worst-case random write scenario:**
  - **Assumptions:**
    - 100% random 8KB transfer writes, 25,000 IOPS
    - Write Amplification of 1.75 with 25% Over-Provisioning
  - **Service Life per GB of NAND:**
    - 8KB * 25,000 * 1.75 = 0.341 GB/s
    - MLC 5K cycles → 5000/0.341 = 4 hours/GB
    - SLC 100K cycles → 100000/0.341 = 81.5 hours/GB
Life Calculation Example

- Worst case sequential write scenario:
  - Assumptions:
    - 100% sequential writes, 500MB/s
    - Write Amplification of 1
  - Service Life per GB of NAND:
    - MLC 5K cycles $\rightarrow$ 5000/0.5 = 2.78 hours/GB
    - SLC 100K cycles $\rightarrow$ 100000/0.5 = 55.56 hours/GB
5 Years Useful Life

- Enterprise workload: 70% read / 30% write
  - Random Writes:
    - MLC / SLC → 13.3 hours/GB / 272 hours/GB
  - Sequential Writes:
    - MLC / SLC → 9.3 hours/GB / 185 hours/GB
- To meet 5 years service life (43830 hours)
  - Random Writes need drive capacity of:
    - MLC → 3.3TB
    - SLC → 161GB
  - Sequential Writes need drive capacity of:
    - MLC → 4.7TB
    - SLC → 237GB
- Monitoring of actual usage is key to predicting remaining service life
## Comparison of Different Methods

<table>
<thead>
<tr>
<th>Family</th>
<th>Capacity</th>
<th>Block Level</th>
<th>LifeEst (W-Yr/GB) (W-Yr)</th>
<th>200GB per Day (Yr)</th>
<th>LDE (TBW) (Yr @ BAPco Pro)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SiliconDrive EP</td>
<td>32GB</td>
<td>100K</td>
<td>0.07 2.24</td>
<td>43.8</td>
<td>3,200 769.2</td>
</tr>
<tr>
<td>SiliconDrive II</td>
<td>32GB</td>
<td>1M</td>
<td>1.7 54.7</td>
<td>438.4</td>
<td>32,000 7692.3</td>
</tr>
<tr>
<td>SiliconDrive 4xxx</td>
<td>4GB</td>
<td>1M</td>
<td>2.7 10.8</td>
<td>54.8</td>
<td>4,000 961.5</td>
</tr>
<tr>
<td>SiliconDrive 3xxx</td>
<td>8GB</td>
<td>2M</td>
<td>10.1 81.1</td>
<td>219.2</td>
<td>16,000 3846.2</td>
</tr>
</tbody>
</table>
Endurance Monitoring

- T13 Device Statistics (Log address 04h, page 07h for SSD)
  - Percent used; other parameters submitted through JEDEC
- Proprietary Implementations
  - SMART (B0h)
  - Vendor specific
Vendor Specific Example

- Use WD SiSMART – over 4M units shipped
  - Load SiliconDrive into the application
  - Run the application for a month or a week
  - Read the SiSMART percentage used data
  - Extrapolate the real world percentage based on the following formulas:

  \[
  \text{Years} = \frac{1 \text{ month}}{\text{SiSMART \%}} \times \frac{1 \text{ year}}{12 \text{ months}} \times 100% \\
  \text{Years} = \frac{1 \text{ week}}{\text{SiSMART \%}} \times \frac{1 \text{ year}}{52 \text{ weeks}} \times 100% 
  \]
SSD Trade-offs

SSD Technology
- Wear-Leveling
- ECC
- NVM Config
- Over Provisioning
- Performance
- Write Amplification

Usage Model
- Storage Budget
- Capacity
- Data Size
- Write Duty Cycle
- Randomness of data

Life (Years) = SSD Technology × Usage Model
SSD Trade-offs

SSD Technology
- Wear-Leveling
- ECC
- NVM Config
- Over Provisioning
- Performance
- Write Amplification
- NVM Specs

Usage Model
- Storage Budget
- Capacity
- Data Size
- Randomness of data
- Write Duty Cycle

Life (Years)

Storage Developer Conference 2009
© 2009 Insert Copyright information here. All rights reserved.
SSD Trade-offs

SSD Technology
- Wear-Leveling
- ECC
- NVM Config
- NVM Specs
- Over Provisioning
- Performance

Usage Model
- Storage Budget
- Capacity
- Data Size
- Randomness of data
- Write Duty Cycle

Life (Years)
Takeaways

- Endurance and performance play against each other
- Workload has the biggest impact on useful life
- NAND media becoming more differentiated
- SSD architectures vary greatly
- For moderate to heavy workloads, SLC provides a better value than MLC
Resources

- Western Digital white paper
  - “NAND Evolution and its Effects on SSD Useable Life”
  - “Storage Design Criteria for Embedded Systems”
- JEDEC – SSD and NAND Media
  - (www.jedec.org)
- SNIA - Storage Networking Industry Association
  - (www.snia.org)
- IDEMA
  - (www.idema.org)
- T13 – Endurance Monitoring
  - (www.t13.org)