Write Amplification due to ECC on Flash Memory or Leave those Bit Errors Alone

Sangwhan Moon and A. L. Narasimha Reddy
Texas A&M University

sangwhan at tamu.edu
reddy at ece.tamu.edu
Introduction (1/2)

- Flash Memory Write Endurance
  - 10,000 P/E cycles for MLC

- Flash Memory Protection Scheme
  - Error Correcting Code (ECC)
  - Scrubbing
  - Wear-leveling and Garbage Collection
  - Parity protection (RAID)

- These protection schemes
  (+) Improve the reliability of flash memory
  (−) Amplify writes → Reduce the reliability of flash memory
Write amplification (W.A.)
- The number of excess writes / writes issued by system

Main sources of W.A.
- Copying live data in garbage collection
- Writing corrected data back in ECC recovery
- Parity update of RAID

W.A. degrades
- performance (related work)
- flash memory’s lifetime
Write Amplification from ECC (1/2)

- W.A. due to ECC recovery
  - Reads lead to writes

```
Read Request

Read a page

ECC check

Error?
  - Yes
    - Correctable? (Yes) => Correct errors
    - Correctable? (No) => Write page
  - No => Read done

Write page

Correct errors

Failure
```
Write Amplification from ECC (2/2)

\[
W.A. = \frac{w'}{w}
\]

Flash device interface

\[
W.A. = \frac{w''}{w}
\]

ECC

\[
w_{ecc} = f_{ecc} \cdot r
\]

Flash device interface

\[
w'' = (w_{ecc} \cdot w')
\]

W.A.

\[
W.A.
\]

Traditional point of view to W.A.

Our point of view to W.A.

→ Severe problem with read intensive workload
Contributions

- A statistical model
  - The impact of the W.A. to the lifetime of flash memory

- A loss of 48% of the lifetime due to the W.A.
  - Various parameters were tested

- Threshold-based ECC to reduce the W.A.
  - Improves the lifetime up to 64%
  - The way to control W.A. to maximize the lifetime
A Reliability Model (1/2)

- **Raw Bit Error Rate from measurement study**

- **Uncorrectable Page Error Rate (UPER)**
  - A Canonical Markov Model

![Diagram](Image)
A Reliability Model (2/2)

- **Uncorrectable Page Group Error Rate (UGER)**

  - Uncorrectable Page Group Error Rate ($\text{UPER}$)
  - Whole Device Failure Rate
  - Page Group Recovery Rate
  - Device Recovery Rate

- **Mean Time to Data Loss**

$$\text{MTTDL}_p = \lim_{k \to \infty} \sum_{j=1}^{k} \left( j g(j) \prod_{i=1}^{j-1} (1 - g(i)) \right)$$

The probability of uncorrectable page group error
Evaluation: Write Amplification

- More read, higher W.A.
- G.C.: -19% lifetime
- ECC: -42% lifetime
- G.C. + ECC: -48% lifetime

<table>
<thead>
<tr>
<th>r:w</th>
<th>5000</th>
<th>10000</th>
<th>15000</th>
<th>20000</th>
<th>25000</th>
<th>30000</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:1</td>
<td>1.0302</td>
<td>1.0839</td>
<td>1.2125</td>
<td>1.4430</td>
<td>1.7011</td>
<td>1.8738</td>
</tr>
<tr>
<td>3:1</td>
<td>1.0308</td>
<td>1.0889</td>
<td>1.2475</td>
<td>1.6287</td>
<td>2.3165</td>
<td>3.0930</td>
</tr>
<tr>
<td>5:1</td>
<td>1.0309</td>
<td>1.0899</td>
<td>1.2560</td>
<td>1.6862</td>
<td>2.5968</td>
<td>3.9032</td>
</tr>
<tr>
<td>7:1</td>
<td>1.0310</td>
<td>1.0904</td>
<td>1.2598</td>
<td>1.7142</td>
<td>2.7571</td>
<td>4.4806</td>
</tr>
<tr>
<td>9:1</td>
<td>1.0310</td>
<td>1.0906</td>
<td>1.2619</td>
<td>1.7308</td>
<td>2.8609</td>
<td>4.9130</td>
</tr>
</tbody>
</table>

W.A. from ECC recovery at different P/E cycles

160GB 3x nm SSD
100MB/s Bandwidth
61bits correctable / 4KB
50% Device Utilization
R:W=3:1
Evaluation: Various Parameters

1. Scrubbing may be harmful (not always)

2. Lifetime highly depends on space utilization

3. Hotness helps to improve lifetime (efficient garbage collection)

4. RAID improves lifetime (Mirroring is the best since it splits read workload)
### Threshold-based ECC (1/5)

- A few bit errors accumulate before ECC correction

<table>
<thead>
<tr>
<th>$n$</th>
<th>5000</th>
<th>10000</th>
<th>15000</th>
<th>20000</th>
<th>25000</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\leq 1$</td>
<td>0.0286</td>
<td>0.0756</td>
<td>0.1657</td>
<td>0.2463</td>
<td>0.2105</td>
</tr>
<tr>
<td>$\leq 3$</td>
<td>0.0295</td>
<td>0.0823</td>
<td>0.2077</td>
<td>0.4022</td>
<td>0.4604</td>
</tr>
<tr>
<td>$\leq 5$</td>
<td>0.0295</td>
<td>0.0824</td>
<td>0.2096</td>
<td>0.4323</td>
<td>0.5824</td>
</tr>
<tr>
<td>$&gt; 5$</td>
<td>6.57e-10</td>
<td>3.12e-7</td>
<td>8.50e-5</td>
<td>0.0072</td>
<td>0.1163</td>
</tr>
</tbody>
</table>

Probability distribution of the number of accumulated bit errors $n$ when they are recovered by ECC.

*58.2% of recoveries for pages with $\leq 5$ bit errors*
Threshold-based ECC (1/5)

- A few bit errors accumulate before ECC correction

### Probability distribution of the number of accumulated bit errors $n$

<table>
<thead>
<tr>
<th>$n$</th>
<th>5000</th>
<th>10000</th>
<th>15000</th>
<th>20000</th>
<th>25000</th>
</tr>
</thead>
<tbody>
<tr>
<td>$= 1$</td>
<td>0.0286</td>
<td>0.0756</td>
<td>0.1657</td>
<td>0.2463</td>
<td>0.2105</td>
</tr>
<tr>
<td>$\leq 3$</td>
<td>0.0295</td>
<td>0.0823</td>
<td>0.2077</td>
<td>0.4022</td>
<td>0.4604</td>
</tr>
<tr>
<td>$\leq 5$</td>
<td>0.0295</td>
<td>0.0824</td>
<td>0.2096</td>
<td>0.4323</td>
<td>0.5824</td>
</tr>
<tr>
<td>$&gt; 5$</td>
<td>6.57e-10</td>
<td>3.12e-7</td>
<td>8.50e-5</td>
<td>0.0072</td>
<td>0.1163</td>
</tr>
</tbody>
</table>

11.6% of recoveries for pages with > 5 bit errors
IDEA: Postpone write until errors accumulate?

Avoid writes until bit errors accumulate to a threshold
Threshold-based ECC (3/5)

- A Markov model for reliability analysis

A Markov model for reliability analysis

- The number of errors in a page
- Page recovery rate
- Garbage collection rate
- Bits/page
- The number of correctable errors
- Bit error rate

\[
\begin{align*}
&S_0(x) \\
&\mu + g(i, x) \\
&(S - 1)\lambda(x) \quad (S - N - 1)\lambda(x) \\
&\quad (S - N)\lambda(x) \\
&\quad (S - K)\lambda(x)
\end{align*}
\]
Threshold-based ECC (4/5)

Evaluation

- Optimal threshold depends on environment and # of devices
- Lifetime improves up to 64%
Threshold-based ECC (5/5)

- Controlling W.A. to achieve max lifetime

Diagram:
- Optimal Recovery Rate
- Lifetime
- Inspection rate
- Maximal Lifetime
- Weak Recovery: insufficient
- Higher W.A.: too often
- Scrubbing
Conclusion

- Reads lead to W.A.
  - A Statistical Reliability Model
  - A loss of 48% of the lifetime due to W.A.

- To control W.A. through two tools
  - Scrubbing for detecting latent errors
  - **Threshold-based ECC** for avoiding excessive recovery
Thank you

Q & A

Sangwhan Moon (sangwhan at tamu.edu)
A. L. Narasimha Reddy (reddy at ece.tamu.edu)