Why the Endurance of cMLC Doesn't Matter

Bill Radke

Director of Architecture, Skyera, Inc

Why the Endurance of cMLC Doesn't Matter

- SLC, eMLC, and cMLC

- Cycling and Endurance

- Mean Time to Data Loss

SLC, eMLC, and cMLC

- ## SLC stores 1 bit per cell
  - ### Lowest ECC requirements
  - ### 100k cycles

- ## Enterprise-MLC stores 2 bits per cell
  - ### Highest ECC requirements
  - ### 30k cycles

- ## Consumer-MLC stores 2 bits per cell
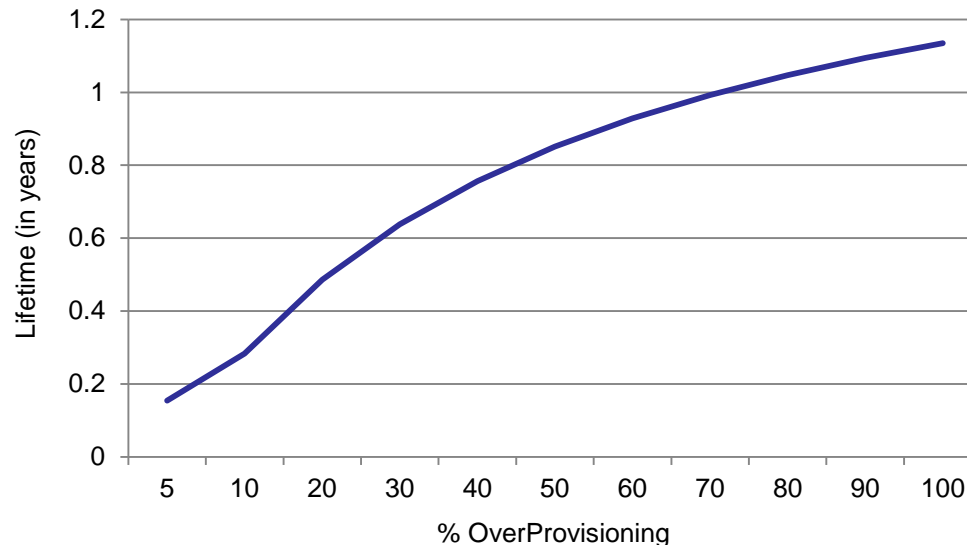 - ### 3k cycles

SLC, eMLC, and cMLC

- **All three parts are the same silicon**
  - Cherry-picking and trimming

- **Adjusted to same ECC failure rate**

- **Measured using an optimized trim set**
 - Everything else at worst case

Cycling and Endurance

- **Superficially, there is a danger to only 3k NAND cycles**

- **Chart below shows lifetime at 1MIOPS**
  - **Parts will show degradation within months**

Cycling and Endurance: Overprovisioning

- **Overprovisioning is simple optimization**
  - It is also very expensive

- **Not necessary for All-Flash Arrays**
  - Application-driven traffic
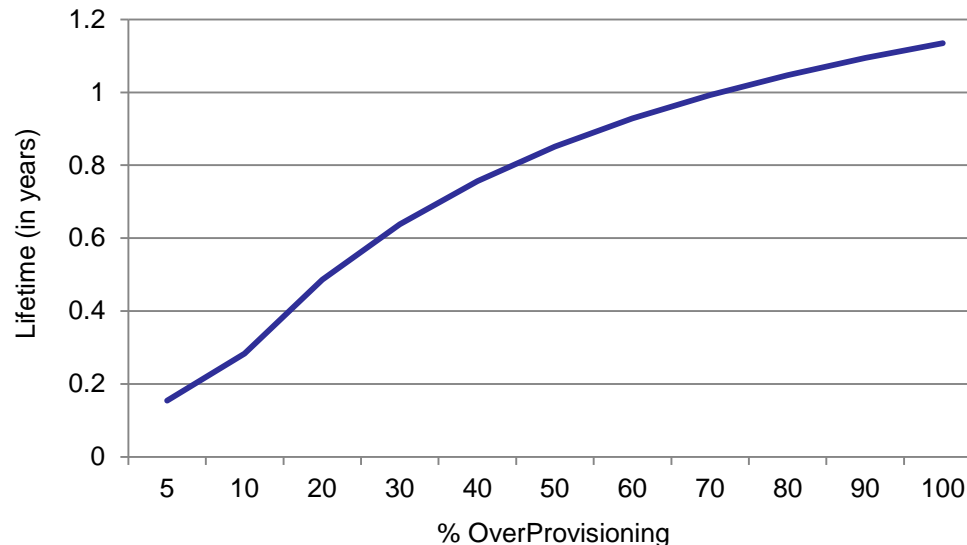
Cycling and Endurance: Alternatives to cMLC

- cMLC and high OP
  - Significantly decreases cycling requirements
- eMLC
  - Higher native capability
- SLC
  - Higher cycling and performance

- cMLC is the lowest cost solution!

Cycling and Endurance: 100x Life Amplification

- By careful, system-wide optimization, the life can be extended by 100x
  - Instead of 0.2 years, 20 years of use
  - Makes cMLC practical

- Combination of decreased Flash traffic and extended Flash life
  - No one feature, instead tight design integration

Cycling and Endurance: Decreased Writes

- ## Inline compression and dedup

- ## Solid-state aware RAID-SE
  - ### Optimal performance for minimal overhead

- ## Garbage collection and scrubbing
  - ### Dynamically optimizes for applications

Cycling and Endurance: Increased Endurance

- **Careful analysis of device physics**
    - Always operating in the sweet spot

- **Adaptive, dynamic retrimming**
    - Optimize NAND settings as needed/able

- **Exact control of operating environment**

Cycling and Endurance: NAND Failures

- Decreased traffic and extended cycling are orthogonal factors

- Keep the ECC failure rate below the spec'd level across the product life

- NAND devices degrade over time
  - Most protection is needed at end-of-life

Mean Time to Data Loss

- **Primary concern is MTTDL**

- **Requires failure of ECC and RAID**
  - CRC & Parity still operating

- **Extremely unlikely within Flash lifetime**
  - Most likely at end of life

Mean Time to Data Loss: ECC Failure

- ## ECC failure is expect to be < 1×10^{-10}
  - ### Spec'd by NAND vendors

- ## Each 10B codewords, expect 1 failure
  - ### Worst-case conditions

- ## At 1M IOPS, a failure every 1.4 hours

Mean Time to Data Loss: RAID-SE Failure

- **Skyera employs a 22+2 RAID-SE arrangement**
  - Two failures during recovery for data loss

- **Odds of first failure is ~ 2.1×10^{-9}**
- **Odds of second failure is ~ 2.0×10^{-9}**

- **Odds of a RAID failure is 4.2×10^{-18}**

Mean Time to Data Loss: Odds to Fail

- Skyera boxes are spec'd for 5 years
  - During the first 4 years, the Flash failure rate is very low

- During the last year, odds of data loss are 2.65×10^{-14}

- The odds of winning the lottery are 6.3×10^{-8}
  - 2M times more likely

Conclusions

- Does the Endurance of cMLC Matter?

Conclusions

- Does the Endurance of cMLC Matter? No!

Endurance of cMLC Doesn't Matter

- **cMLC has enough cycles for a 5 year life**
  - **If the cycles are carefully utilized!**

- **Data with errors can be recovered**
  - **Even at the end of life**

- **All this requires a system-wide approach**