Lifetime Improvement of NAND Flash-based Storage Systems Using Dynamic Program and Erase Scaling

FAST’14
Contents

• Lifetime Issues
• Motivations
• Key Components of DPES
• Implementation of DPES-aware FTL
• Experiments and Conclusions
NAND Device Trend

- NAND storage capacity is increasing
  - Semiconductor process scales to 10nm
- Erase-before-program property
  - Repeated P/Es make NAND cell wear down
  - NAND device provides limited number of P/E cycles
- Thin paper (highly scaled device) wears down more easily
  - NAND lifetime (endurance, P/E cycle) becomes a main barrier
Directions to Improve Lifetime

\[ L_c = \frac{\text{MAX}_{P/E} \times C}{\text{W}_{\text{day}} \times \text{WAF}} \]

- Lifetime (in days) of NAND device, \( L_c \)
  - (+) Maximum number of P/E cycles, \( \text{MAX}_{P/E} \)
  - (+) Total capacity, \( C \)
  - (-) Total written data, \( \text{W}_{\text{day}} \)
  - (-) Write amplification factor, \( \text{WAF} \)
- Reducing WAF (by increasing FTL efficiency)
- Reducing \( \text{W}_{\text{day}} \) (by deduplication, compression)
- Increasing \( \text{MAX}_{P/E} \) (by self-recovery)
- DPES tries to increase \( \text{MAX}_{P/E} \)
Motivations: Device Physics

1. Using softer eraser (using low erase voltage)
   - Paper wears down slowly
   - NAND cells wears out slowly
   - Device endurance (lifetime) can be improved

2. Erasing carefully (increasing erase time)
   - Same effect as using soft eraser
Overview of DPES

DPES (Dynamic Program and Erase Scaling) approach:

- Erase Voltage Scaling
  - Endurance vs. Erase Voltage
- Erase Time Scaling
  - Endurance vs. Erase Time
- Program Time Scaling
  - Erase Voltage vs. Program Time

Dynamically changes program and erase voltage/time

Improves the NAND endurance w/o degradation in the overall write throughput

- (1) Erase voltage scaling (using soft eraser)
- (2) Erase time scaling (erasing carefully)
- (3) Program time scaling (for resolving the side-effect of erase voltage scaling)
- DPES dynamically changes (1)~(3) at runtime
Erase Voltage Scaling

• **BER over different P/E cycles under varying erase voltage**
  – r=0 (nominal vol.), r=0.07 (7% reduced vol.), r=0.14 (14% reduce~)
  – Bit Error Rate (BER) decreases when erase voltage reduced

• New metric: **effective wearing (per P/E), EW**
  – Normalized BER after 3K P/E cycles using nominal voltage

• **Effective wearing over different erase voltage scaling**
  – EW near-linearly decreases as voltage decreases
  – 14% erase voltage reduction → 54% increase of EW
• 2-bit MLC uses four distinguished **threshold voltage levels**
  – Four stages are distinguished by **reference voltages**
  – **Write** injects electrons to floating gate by the corresponding threshold voltage level
  – **Read** checks the current threshold voltage level
• Incremental step pulse programming (ISPP) for write
  – Gradually increases write voltage by $V_{ISPP}$ and checks it
  – # of ISPP loops determines write time ($T_{prog}$)
  – $V_{ISPP}$ determines # of ISPP
• Width of threshold voltage distribution
  – Fixed in the design/manufacture time
  – If large, VISPP can be bigger
  – If small, VISPP should be small
• Small VISPP value (fine-grained control)
  – Increases # of ISPP loop
  – Increases write time
Side Effect of Erase Voltage

- Decrease in erase voltage for minimizing effective wearing
- Decrease in threshold voltage window (shallowly erased)
- Narrow distribution of remaining 3 threshold voltage state
- Fine-grained write is required for shallowly erased block
- **Write time will increase** for those blocks
Trade off: Erase Voltage &

- 5 different voltage erase mode
  - Evmode(0) – nominal erase (deep/full/normal erase)
  - Evmode(4) – weakest ease (most shallow erase)
- **Write time** depends on *how much shallowly erased*
  - Program time scaling is needed for erase voltage scaling
- Corresponding write mode (different write time) provided
  - Wmode(0) – nominal write, Wmode(4) takes twice more time
Erase Time Scaling

- As erase time increases, effective wearing decreases
  - 300% increased erase time reduces EW by 19%
- In addition to low voltage erase, slow erase further increases the endurance
- Two modes (ESmode\text{fast} and ESmodeslow) are supported
Endurance Model for DPES

- For improving endurance (decreasing effective wearing)
  - Five erase voltage mode (erase voltage scaling)
  - Two erase time mode (erase time scaling)
- For shallowly erased blocks,
  - Five program time mode (program time scaling)
- DPES dynamically adjust voltage and time mode to minimize effective wearing (improving lifetime) without write time loss
Overview of DPES-enabled FTL (autoFTL)

- **Mode selector** determines write/erase modes
- **Per-Block Mode Table** groups blocks by five erased levels
- **Device Settings** (write, erase, reference voltage) configured
Write/Erase Mode Selection Flow

Write Request
- Circular Buffer
  - Utilization
    - Mode Selector
      - Wmode (i)
      - EVmode (j)
      - ESmode (k)
    - NAND Setting Table
      - DeviceSettings
      - NAND Flash Memory

Write Request
- Select Wmode (i)
- GC is needed?
  - no
  - Select EVmode (j)
  - Select ESmode (k)
  - Erase (j, k)
  - Write (i)
  - no
Write Mode Selection

- Write requests queued into buffer before service
- Buffer utilization (how much the buffer is filled)
  - High utilization $\rightarrow$ intensive writes $\rightarrow$ fast write
  - Low utilization $\rightarrow$ idle time between writes $\rightarrow$ slow write
- Utilization $< 20$
  - Wmode4 is set (if current write mode is different)
  - Write request in the head of buffer is programmed
Erase Voltage/Time Mod

- Note that erase is making free space for future writes
- Voltage erase should consider future write patterns
  - If too many blocks erased with Evmode(4)
  - If future writes are intensive
    - Write time is delayed (using only Wmode(4) is possible)
- When a write, (if no GC), set Wmode and program
- When a write, (if GC), it estimates future buffer utilization, set the Evmode, and erase victim blocks
Experimental Setup

- NAND configuration
  - 8 channels * 4 chips/channel = 32 chips
- Write buffer size
  - 32MB = 8KB page * 4096 write requests
- 6 enterprise traces
  - Based on inter-arrival time ratio, the effectiveness of DPES varies
Endurance Gain

- On average, 69% endurance increased by DPES
- Src1_2 shows less endurance gain
  - It has a large amount of small inter-arrival time requests
  - Writes with EVmode0 are many
- Prxy_0 shows the largest endurance gain
  - Majority writes use EVmode4
Conclusions

• Device physics
  – Erase with low voltage → endurance improved
  – Slow erase → endurance improved
  – Side effect: erase with low voltage → program time scaling

• DPES (Dynamic Program and Erase Scaling)
  – 5 erase voltage/program time mode, 2 erase time mode
  – Endurance model used

• DPES enabled FTL
  – Write time mode selection based on inter-arrival time
  – Erase voltage/time mode selection based on guessing future write inter-arrival time