Optimizing NAND Flash-Based SSDs via *Retention Relaxation*

Ren-Shuo Liu*, Chia-Lin Yang*, Wei Wu†

*National Taiwan University and †Intel Corporation

Motivation

- **Retention specification vs. actual retention requirement**

  Industrial standards: 1 to 10 years **vs.** Applications' needs: days or shorter

- **Retention relaxation**
  - If we don't need to guarantee the maximal retention time, which design parameters of SSDs can be improved?
- **Contribution**
  - We propose retention-aware designs to trade data retention for the benefits on write speed, or ECCs' cost and performance

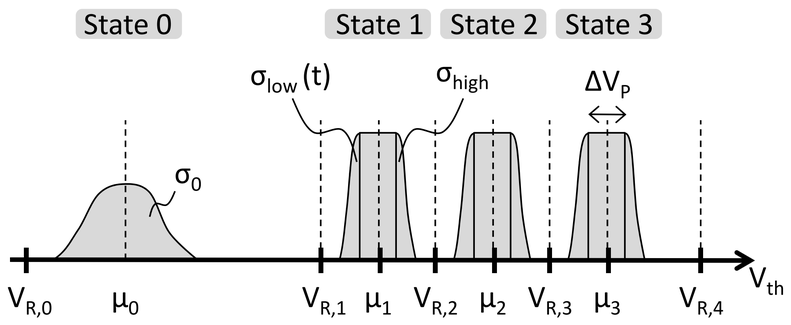
NAND Flash Model

- **Threshold voltage (Vth) distribution model**
  - Probability density function of cells' V_{th}

$$P_k(v,t) = \alpha_0 \cdot e^{-\frac{(v-\mu_0)^2}{2\sigma_0^2}} \quad , k = 0$$

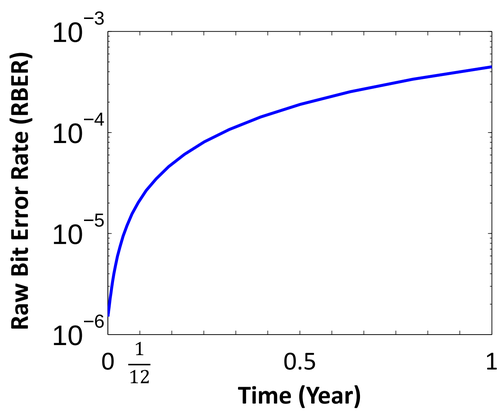
$$P_k(v,t) = \begin{cases} \alpha(t) \cdot e^{-\frac{(v-\mu_k+0.5\Delta V_P)^2}{2\sigma_{low}(t)^2}} & , \ v < \mu_k - \frac{\Delta V_P}{2} \\ \alpha(t) \cdot e^{-\frac{(v-\mu_k-0.5\Delta V_P)^2}{2\sigma_{high}^2}} & , \ v > \mu_k + \frac{\Delta V_P}{2} \\ \alpha(t) & , \ otherwise \end{cases} \quad , k \neq 0$$

where k is the data state #, and ΔV_P is the step increment in NAND Flash's programming procedure

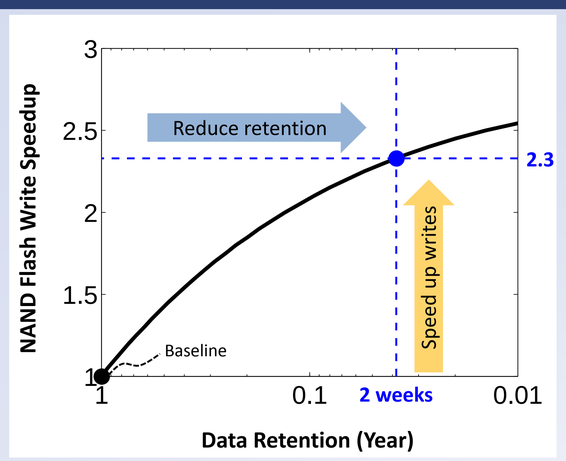


- **Raw bit error rate (RBER)**

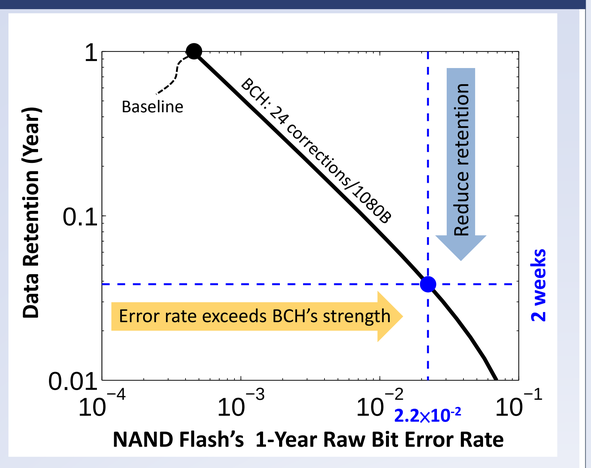
$$RBER(t) = \sum_{k=0}^{q-1} \left(\underbrace{\int_{-\infty}^{V_{R,k}} P_k(v,t)dv}_{V_{th}\text{lower than intended}} + \underbrace{\int_{V_{R,(k+1)}}^{\infty} P_k(v,t)dv}_{V_{th}\text{higher than intended}} \right)$$



Benefits of Retention Relaxation



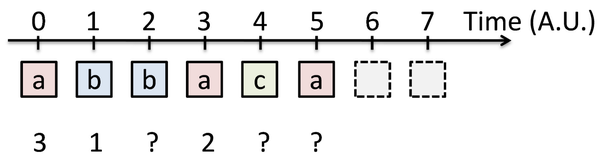
- **Improving write speed**
  - Enlarging the step increment (ΔV_p) in NAND Flash's programming procedures
    - Retention decreases because the V_{th} distributions become wider and the margins between neighboring levels get narrower
    - Write speed increases because with large ΔV_p, fewer programming steps are required during writes
  - 2.3× write speedup is achievable if data retention is reduced to 2 weeks



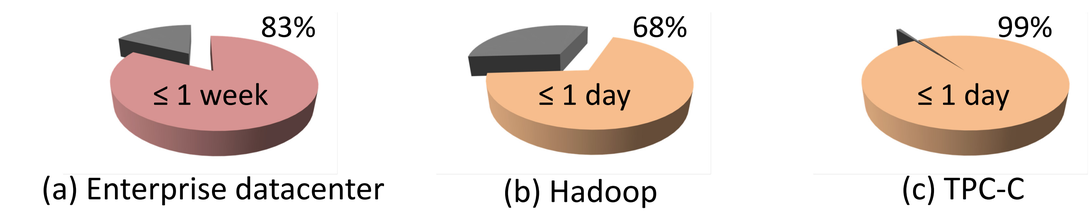
- **Improving ECCs' cost and performance**
  - Advanced ECCs such as LDPC codes are required for NAND Flash whose bit error rate ≥ 10^{-3}
  - Shorter retention guarantee → fewer retention errors → less required ECC strength
  - BCH is strong enough for NAND Flash with the bit error rate up to 2.2×10^{-2} if the retention guarantee is relaxed to 2 weeks

Retention Time Analysis

- **Retention requirement of a sector written into disks**
  - Interval from the time the sector is written to the time the sector is overwritten

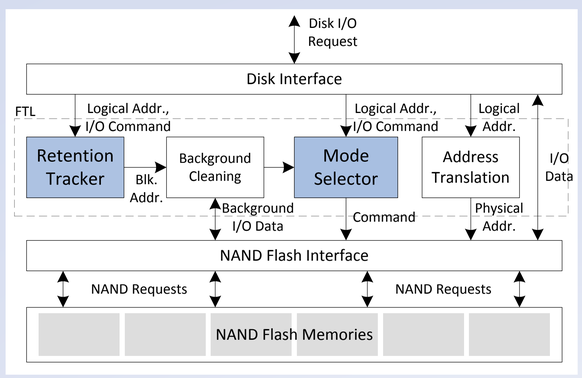


- **68% to 99% of writes have retention requirements ≤ 1 week**

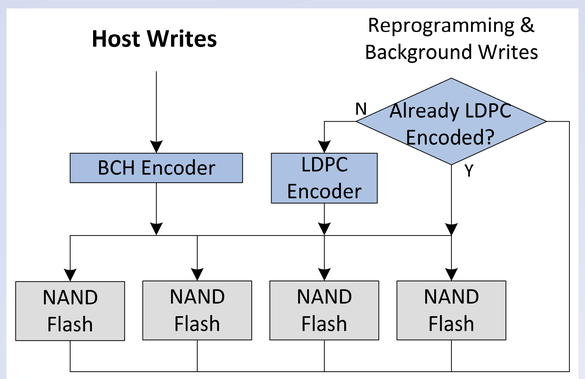


(a) Enterprise datacenter (b) Hadoop (c) TPC-C

Retention-Aware Design

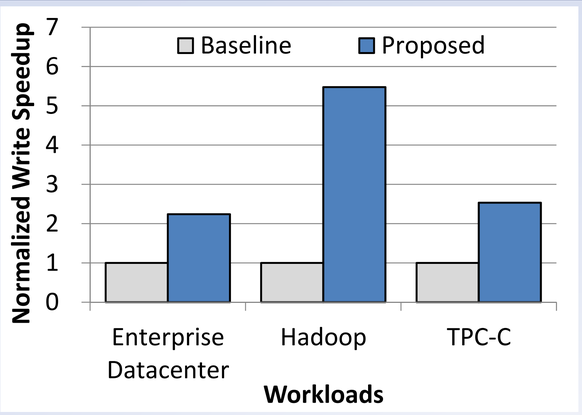


- **Retention-aware Flash Translation Layer (FTL)**
  - Two new components in the FTL
    - **Mode Selector**: invoke different NAND Flash write commands or different ECC engines
    - **Retention Tracker**: monitor blocks and reprogram data which are about to run out of retention to ensure no data loss
  - Two-level retention guarantees
    - **Relaxed retention**: all writes from hosts to SSDs
    - **Normal retention**: background writes (e.g., cleaning and wear-leveling in SSDs) and reprogramming

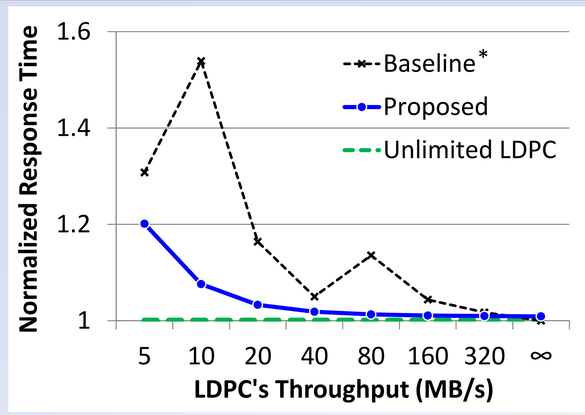


- **Retention-aware ECC architecture**
  - All writes from hosts to SSDs
    - Protected by BCH only
    - Short data retention guarantee
  - Background writes and reprogramming
    - Protected by LDPC
    - Full data retention guarantee
  - Advantages
    - Time-consuming LDPC is kept out of the critical performance path
    - LDPC encodes only data with retention longer than BCH's guarantee

System Evaluation



- **SSD write response time speedup**
  - Retention Relaxation achieves 2.2× to 5.5× speedup
  - Hadoop has the largest speedup
    - High I/O throughput and long queuing time
    - Retention Relaxation significantly reduces the queuing time



- **SSD performance vs. various LDPC throughput**
  - Retention Relaxation outperforms the baseline (conventional concatenated BCH-LDPC) with the same LDPC throughput
  - Retention Relaxation approaches the ideal performance with 20MB/s LDPC

* The SSD simulator could stop simulations due to I/O queue saturation if LDPC's throughput is insufficient. The curve of the baseline presents a zigzag appearance between 5—80 MB/s because several traces which cause saturation are excluded.