Relieving Self-Healing SSDs of Heal Storms

Li-Pin Chang, Sheng-Min Huang, Kun-Lin Chou
Speaker : Sheng-Min Huang

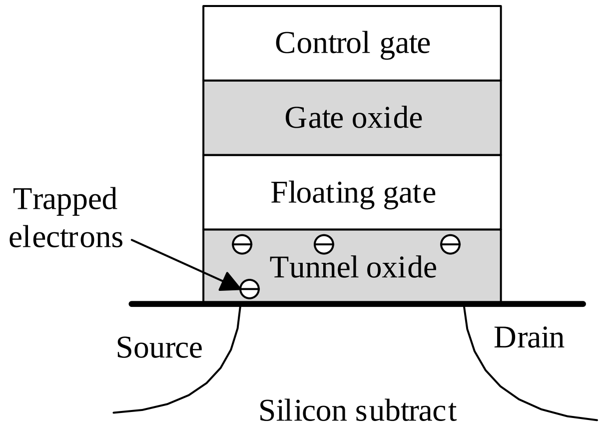
Embedded Software and Storage Lab
National Chiao-Tung University, Taiwan

Outline

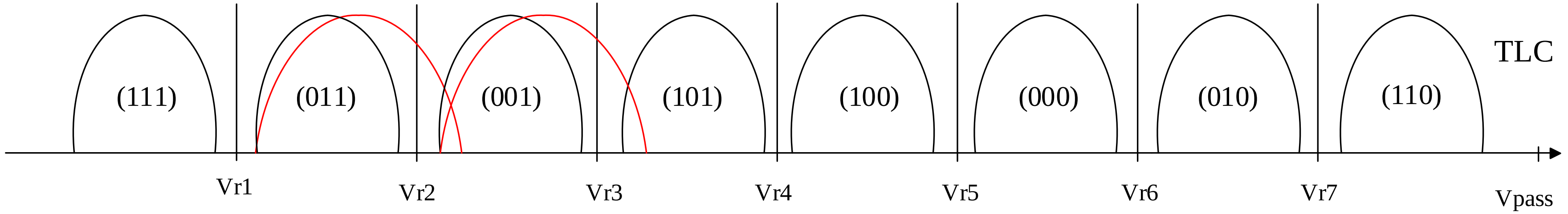
- Introduction
- Heal Storm
- Virtual Wear Leveling
- Experiment Results
- Conclusion

Flash Wear Out Dynamics

- Charge may trap in tunnel oxide after PE cycles.

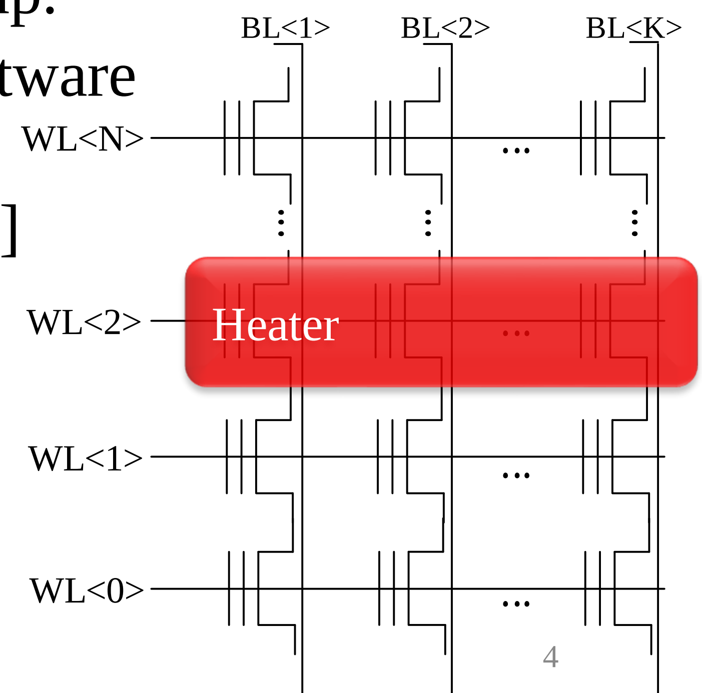


- The threshold voltage shift will become intolerably large and create erroneous bit values.



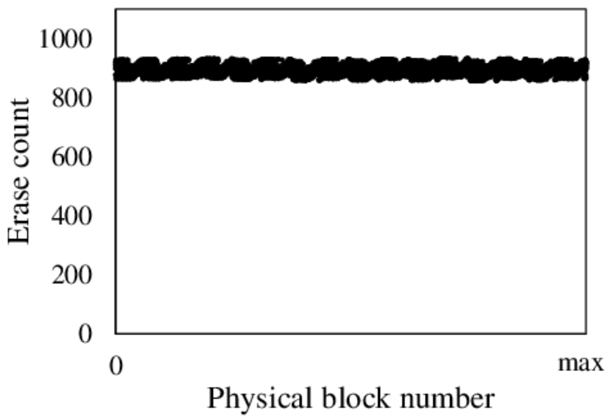
Flash Healing

- Trapped charge (stress) dissipate slowly over time.
  - Accelerate this process under high temperature.
- Healing : Heated-Accelerated Self Recovery [2, 3]
  - Word line heaters to create high temp.
  - Block heal operation for system software
    - Heal nearly worn-out blocks
  - Time cost is about **one second** [3, 4]
  - Relieve about 80% stress [5]

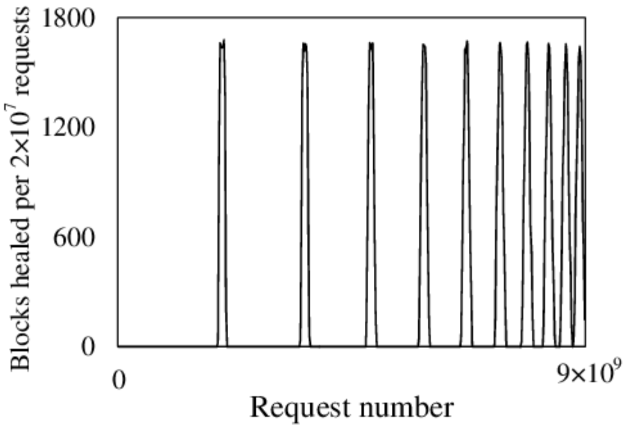


Heal Storms

- Wear Leveling (WL)
  - Strive to balance the erase count of all blocks
- Self-healing flash memory heals flash blocks when blocks reach their PE cycle limit.
- **Heal storm**, blocks undergo block-healing within a short period of time.



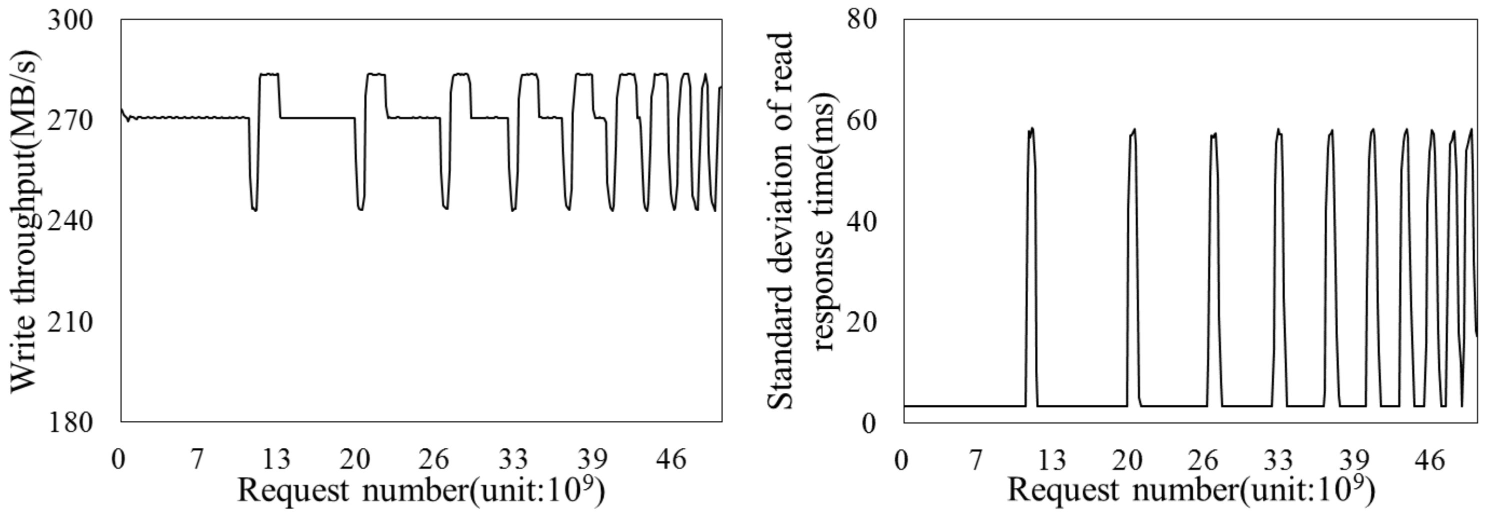
(a)



(b)

Negative Effects of Heal Storms

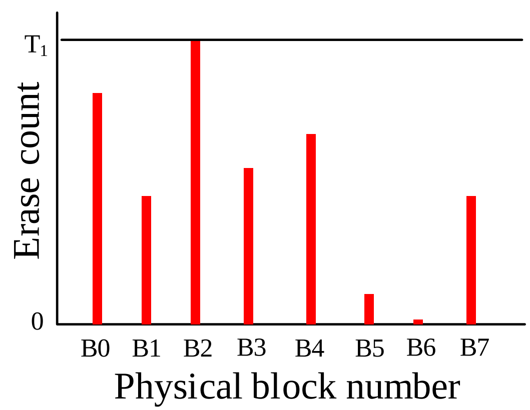
- Read Response degradation
- Write throughput fluctuation
- Unpredictable reliability



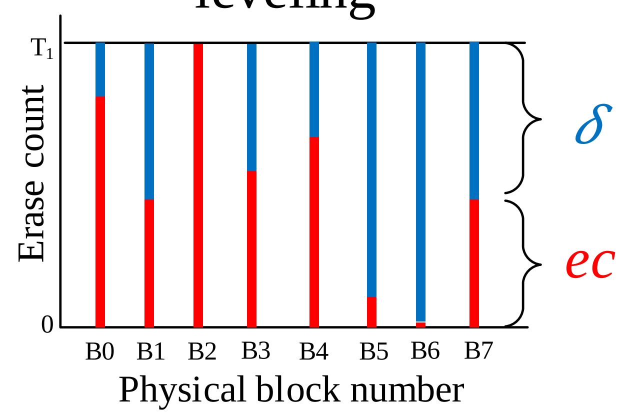
Virtual Wear Leveling

- Leverage the effect of erase count balancing from WL
- Virtual erase count
  - $vec_i = eci + \delta_i$
  - Operate conventional WL on vec

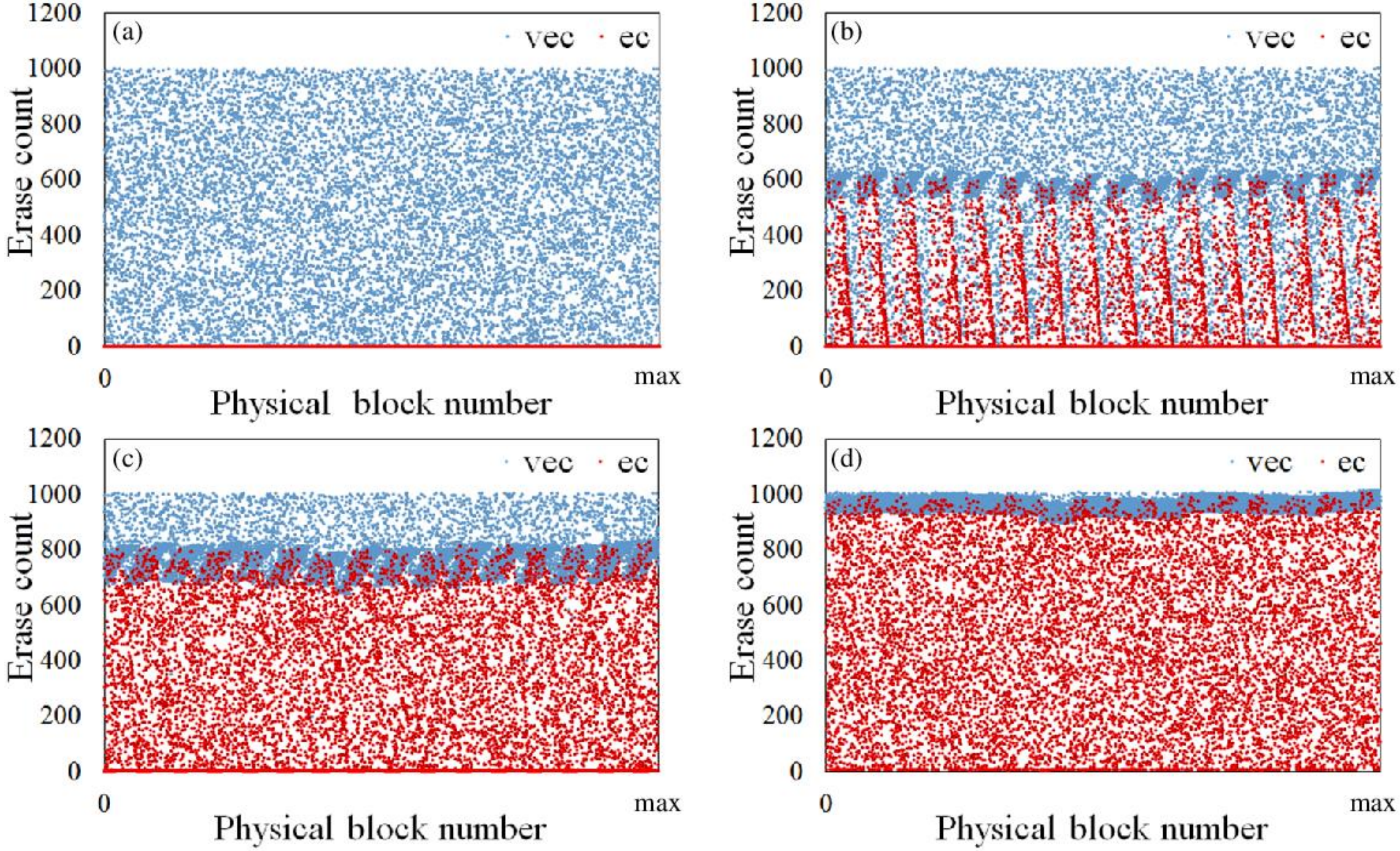
Desired effect



Leveraging wear leveling

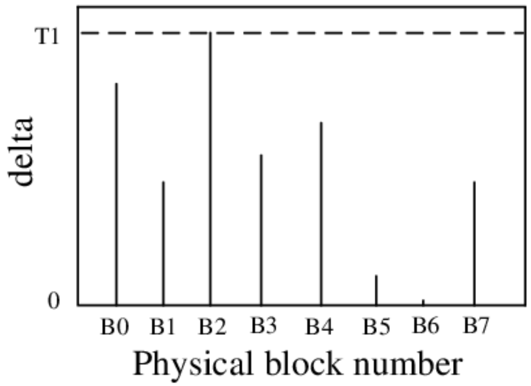


Virtual Wear Leveling in Action

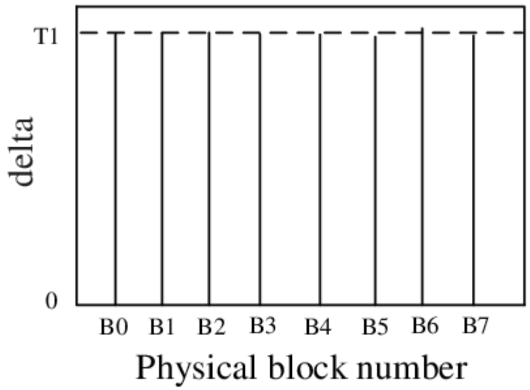


Progressive Delta Leveling

- In the rest of the SSD lifetime,
  - the difference among erase counts remains unchanged.
    ⇒ Lots of blocks have unused PE cycles.
  - all blocks have the same δ (i.e. the difference = 0)
    - Increase δ_i with different rate
    - Update δ_i only after block healing



(a)



(b)

Experiment Setup

- Flash memory parameters
  - 16 flash chips
  - 16 KB per page
  - 4MB per block
  - Latency
    - 0.5 ms for page read
    - 1.6 ms for page write
    - 2.9 ms for block erase
    - 1024 ms for block heal

Workload	Disk vol. size	Total write	Total read	Avg req. size write	Avg req. size read	Scale factor
HM0	39	92	49	37.8	36.3	1/1650
RSRCH0	277	46	5	37.7	38.9	1/4000
STG0	113	62	15	38.1	52.8	1/2700
STG1	113	27	117	36	88.3	1/1900
WDEV0	136	31	9	35.8	40.7	1/5200
EX9	400	220	52	47.1	92.2	1/180
MSN	169	96	321	33.6	35	1/20
	GB	GB	GB	KB	KB	

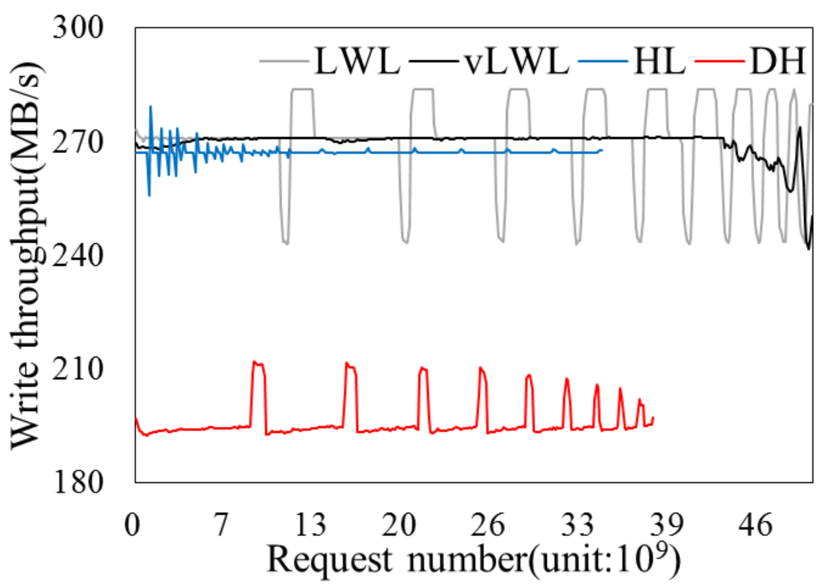
Table 1: Experimental workloads and their characteristics.

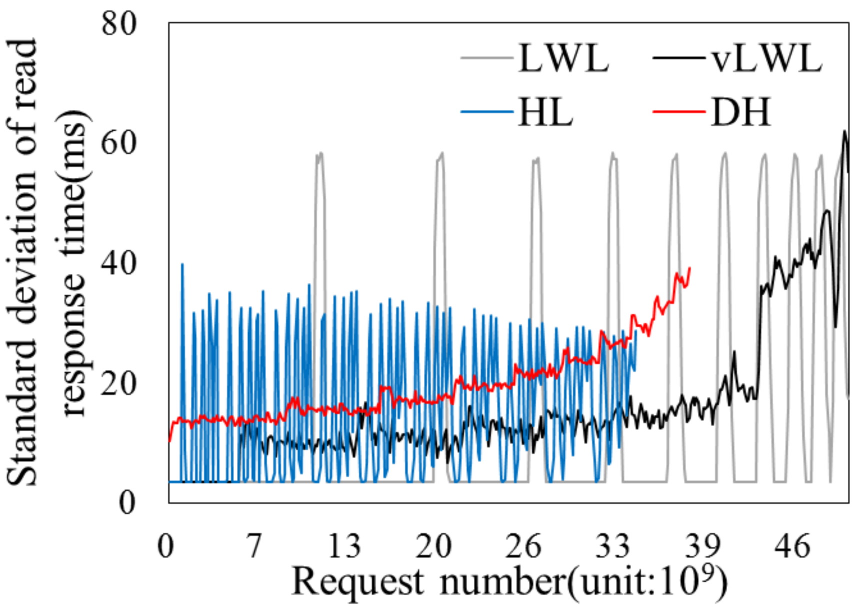
Name	Mapping scheme	Wear Leveling algorithm	Virtual wear leveling
LWL	Page mapping	Lazy Wear Leveling	No
vLWL	Page mapping	Lazy Wear Leveling	**Yes**
HL	Page mapping	Heal Leveling	No
DH	Page mapping	Dheating	No

Table 2: Flash management methods involved in our experiments.

I/O Performance

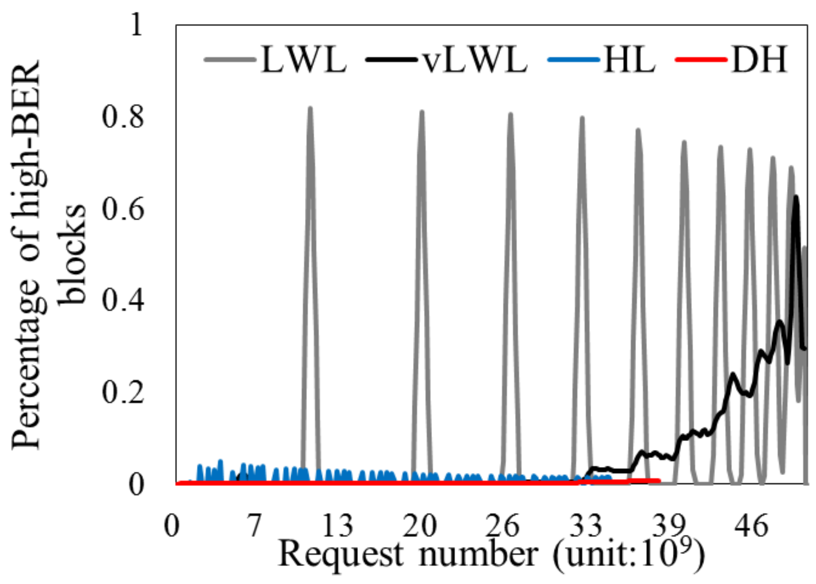
- LWL suffered transient variation by heal storms.
- DH had low write throughput because of high garbage collection overhead.





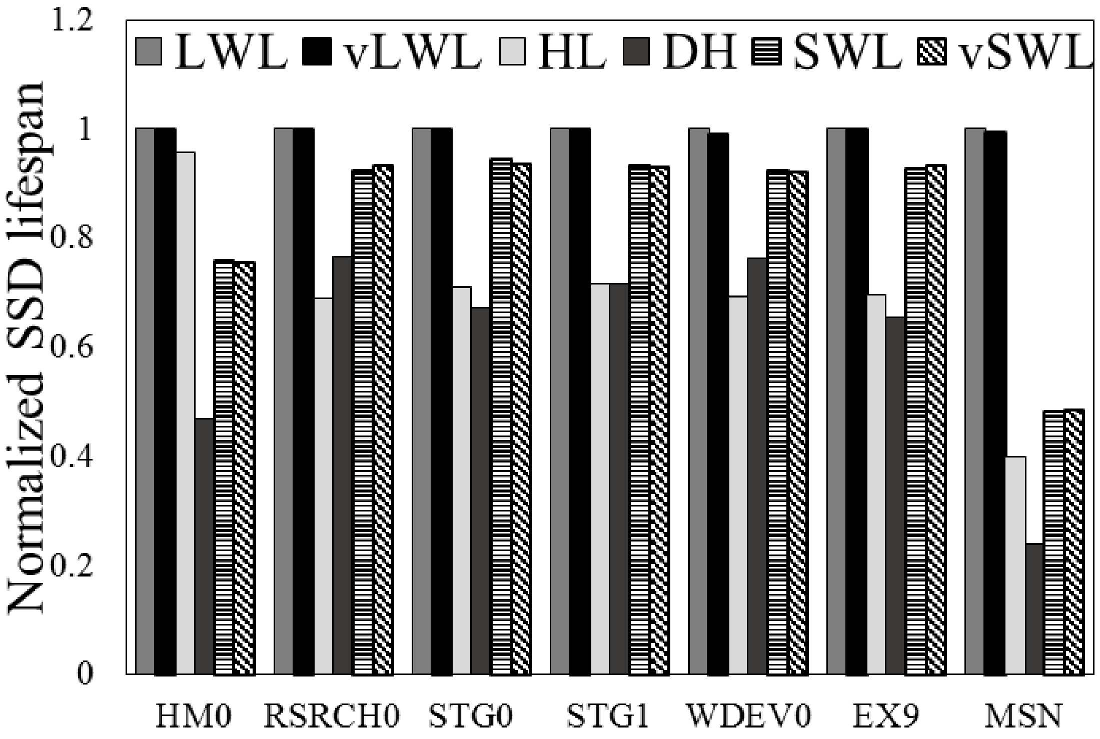
Reliability

- The % of blocks with a high bit error rate (BER) [6] should not fluctuate over time.
- Increasing gradually was good for system software to predict the SSD retirement.



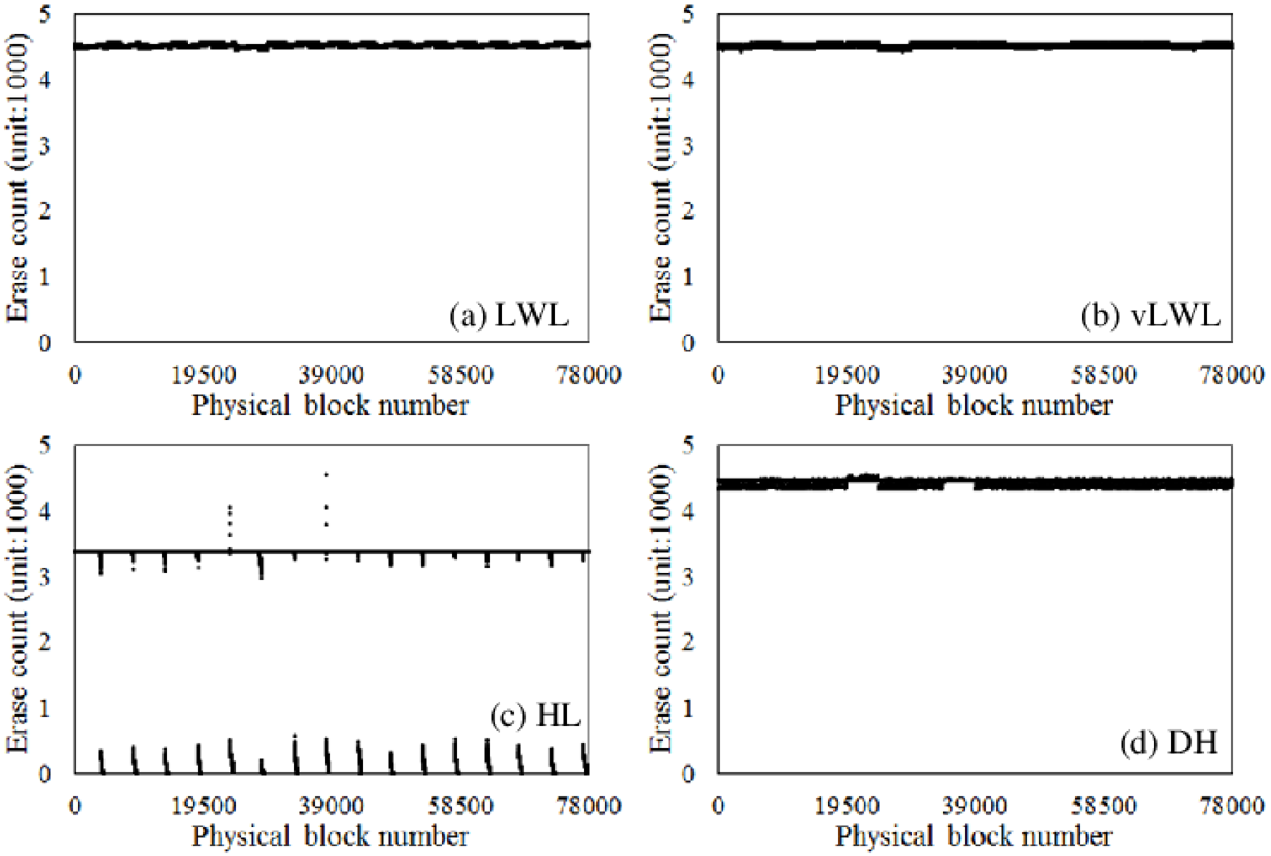
Lifespan

- Our method did not affect the SSD lifespan



Erase count distribution

- Many PE cycles in blocks were wasted in HL.



Experimental Result Summary

- Conventional Wear leveling
  - Suffered heal storms that flash memory were occupied by block-healing operations.
- Dheating
  - Extremely high write amplification because of inaccurate hot/cold identification and local garbage collection in pools.
- Heal Leveling
  - Unexpected short device lifespan because of large variation in erase counts.

Conclusion

- Software-controlled block healing radically extends the SSD lifespan.
- Heal storm damages predictability of performance and reliability.
- Virtual wear leveling leverages conventional wear leveling to disperses block healing over time.
- Possible application of virtual wear leveling
  - Software-controlled bit density [7]
  - Erasing in MLC mode: vec+=2.2
  - Erasing in SLC mode: vec+=1

Thank you

Q & A

Reference

- [1] K. Votto. Samsung SSD 840: Testing the endurance of TLC NAND, 2012.
- [2] Y.-T. Chiu. Forever flash. Spectrum, IEEE, 49(12):11–12, 2012.
- [3] H.-T. Lue, P.-Y. Du, C.-P. Chen, W.-C. Chen, C.-C. Hsieh, Y.-H. Hsiao, Y.-H. Shih, and C.-Y. Lu. Radically extending the cycling endurance of flash memory (to> 100m cycles) by using built-in thermal annealing to self-heal the stress-induced damage. In Electron Devices Meeting (IEDM), 2012 IEEE International, pages 9.1.1–9.1.4, 2012.
- [4] Y.-M. Chang, Y.-H. Chang, J.-J. Chen, T.-W. Kuo, H.-P. Li, and H.-T. Lue. On trading wear-leveling with heal-leveling. In Proceedings of the 51st Annual Design Automation Conference, DAC ’14, 2014.
- [5] Q. Wu, G. Dong, and T. Zhang. Exploiting heat-accelerated flash memory wear-out recovery to enable self-healing SSDs. In USENIX Workshop on Hot Topics in Storage and File Systems (HotStorage), 2011.
- [6] L. Shi, K. Qiu, M. Zhao, and C. J. Xue. Error model guided joint performance and endurance optimization for flash memory. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 33(3):343–355, March 2014.
- [7] X. Jimenez, D. Novo, and P. Ienne. Software controlled cell bit-density to improve NAND flash lifetime. In Proceedings of the 49th Annual Design Automation Conference, DAC ’12, pages 229–234, New York, NY, USA, 2012. ACM.