

Temperature-Aware Data Allocation Strategy for 3D Charge-Trap Flash Memory

Yi Wang¹, Mingxu Zhang¹, Jing Yang²



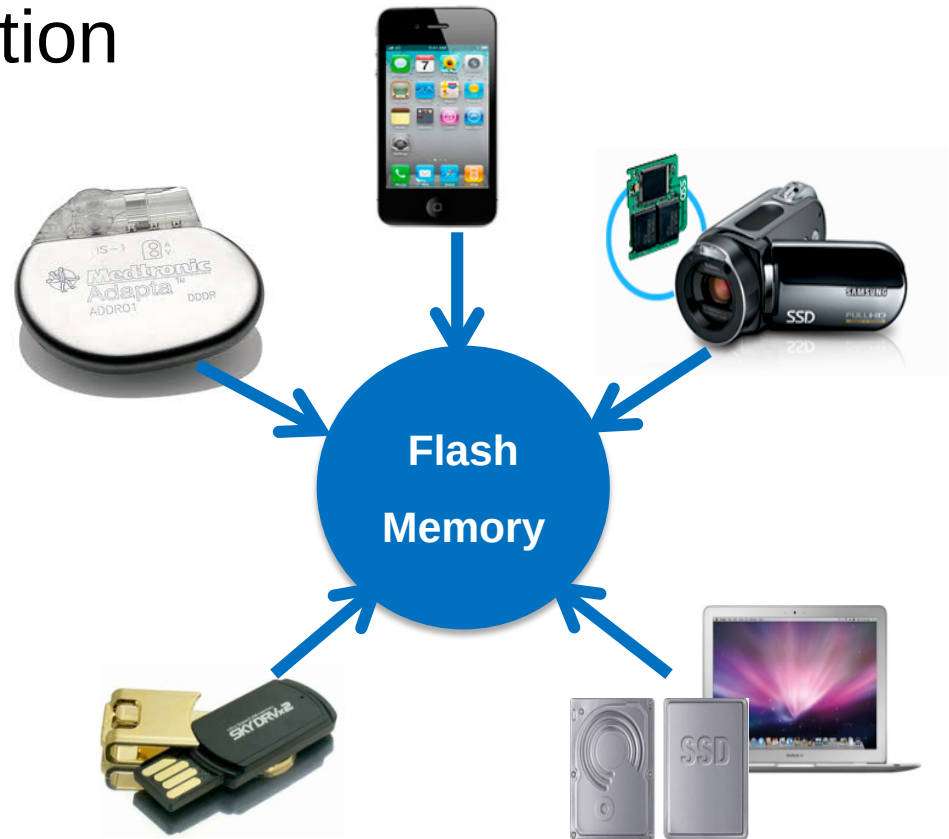
¹ Shenzhen University

² Harbin Institute of Technology
yiwang@szu.edu.cn



Flash Memory Properties

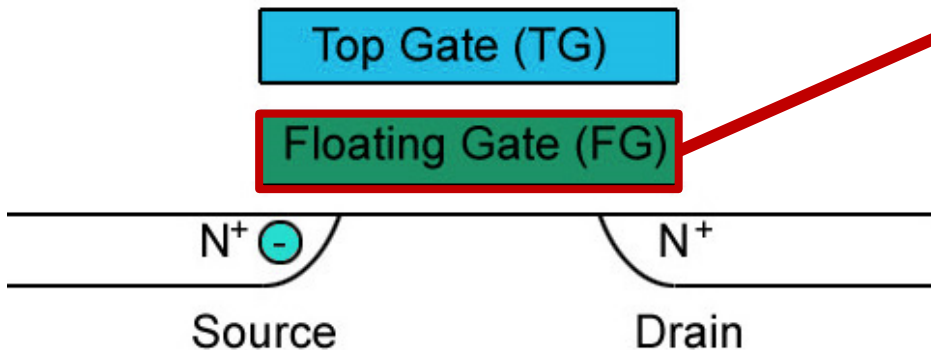
- Faster access performance
- Lower power consumption
- Smaller size
- Lighter weight
- Shock resistance



2D (Planar) Flash Memory

- Floating Gate (FG)

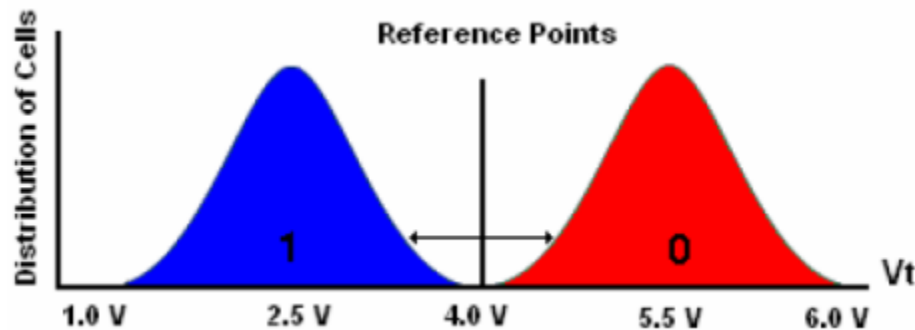
NAND Memory Cell



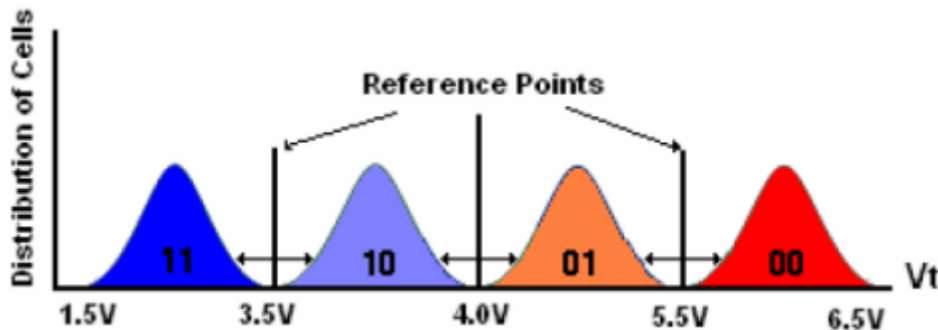
The number of electrons on the floating gate affects the threshold voltage of the cell.

SLC & MLC

- Both SLC and MLC use the threshold voltage to manipulate the state of the flash



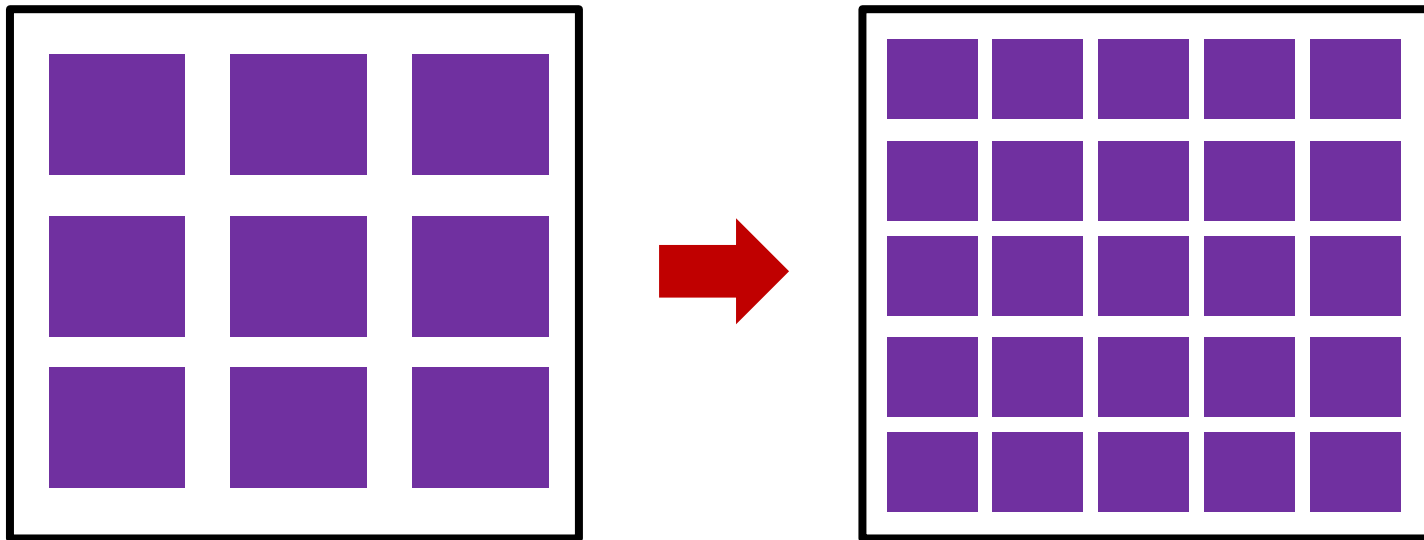
Value	State
0	Programmed
1	Erased



Value	State
00	Fully Programmed
01	Partially Programmed
10	Partially Erased
11	Fully Erased

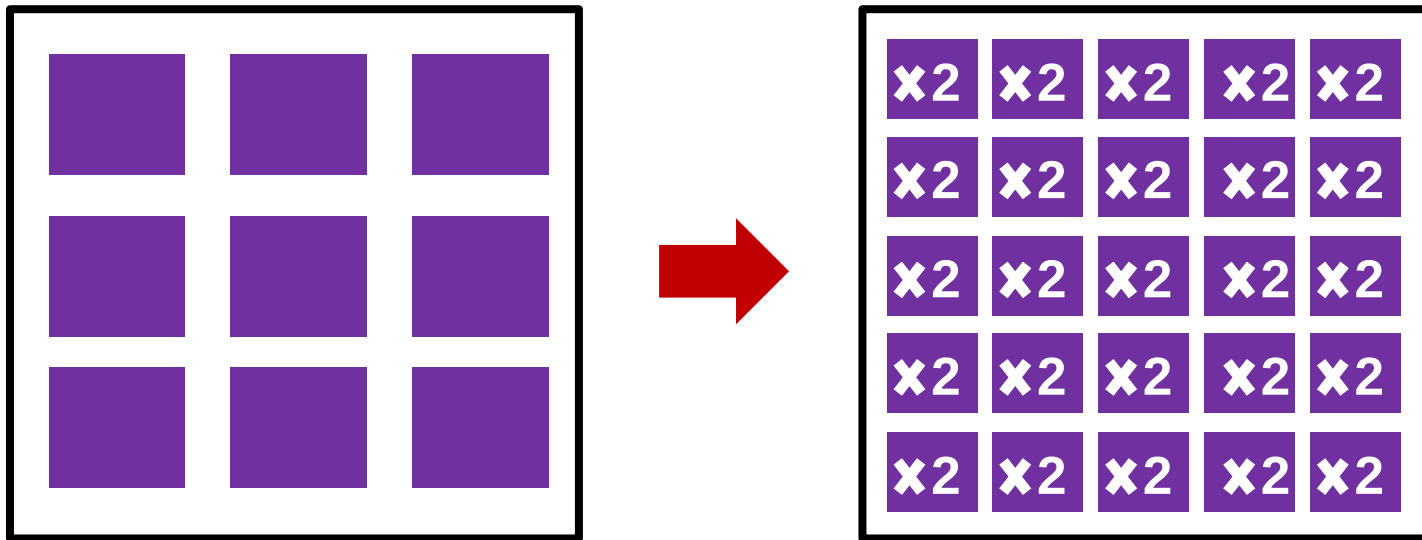
2D Flash → 3D Flash

- Process Improvement



2D Flash → 3D Flash

- Process Improvement
- SLC → MLC → TLC → QLC



2D Flash → 3D Flash

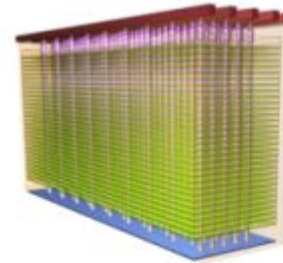
- Floating Gate (2D) → Charge-Trap (3D)



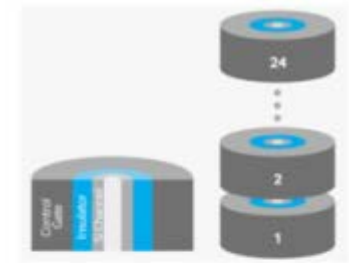
3D Flash Memory



TOSHIBA / Sandisk



Intel / Micron

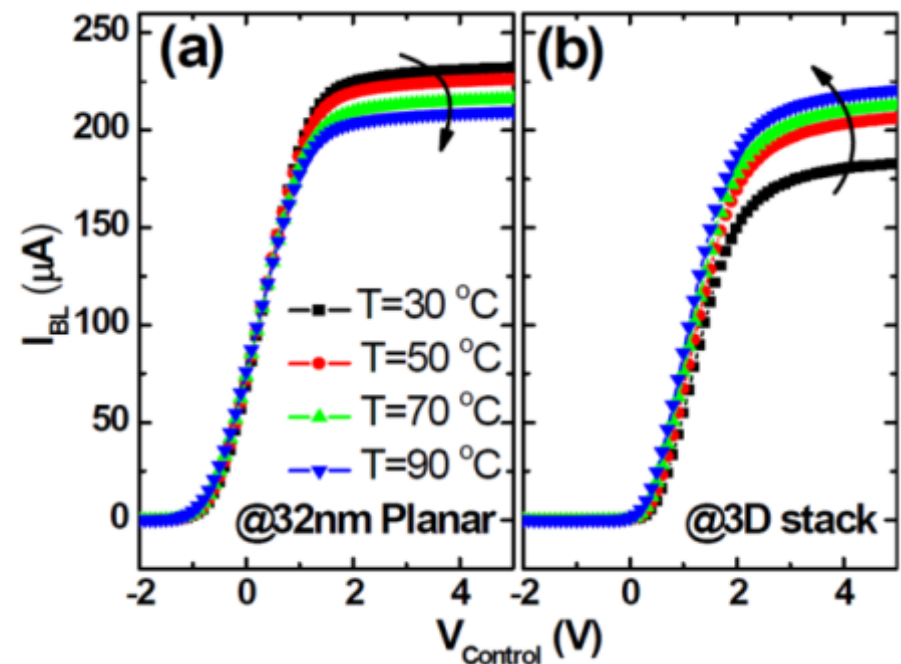
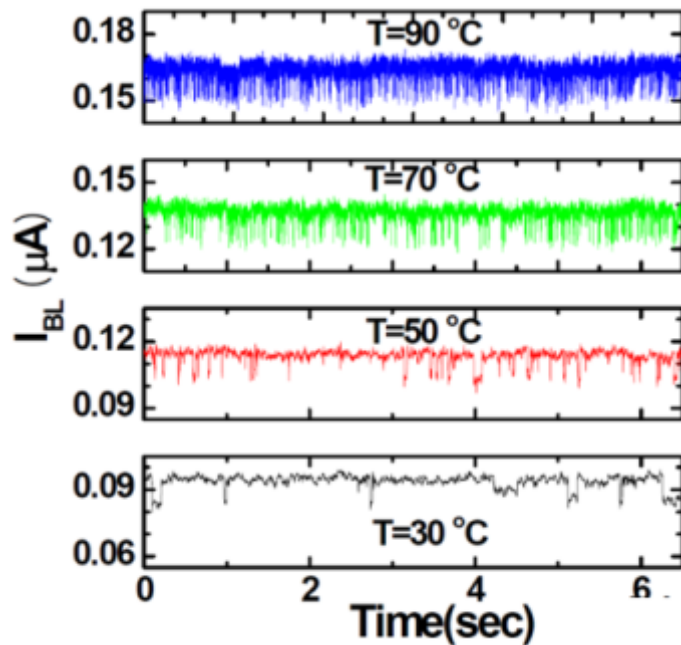


SAMSUNG

Technology	BiCS	FG	V-NAND
Layers	48	32	48
Release	2015	2015	2013
Capacity	128Gb (MLC) 256Gb (TLC)	256Gb (MLC) 384Gb (TLC)	128Gb (MLC) 256Gb (TLC)

Thermal is a critical issue for 3D flash

- The variation of I_{BL} waveform is strongly activated by the temperature [1-2].
- High temperature introduces charge loss and worse retention behavior.



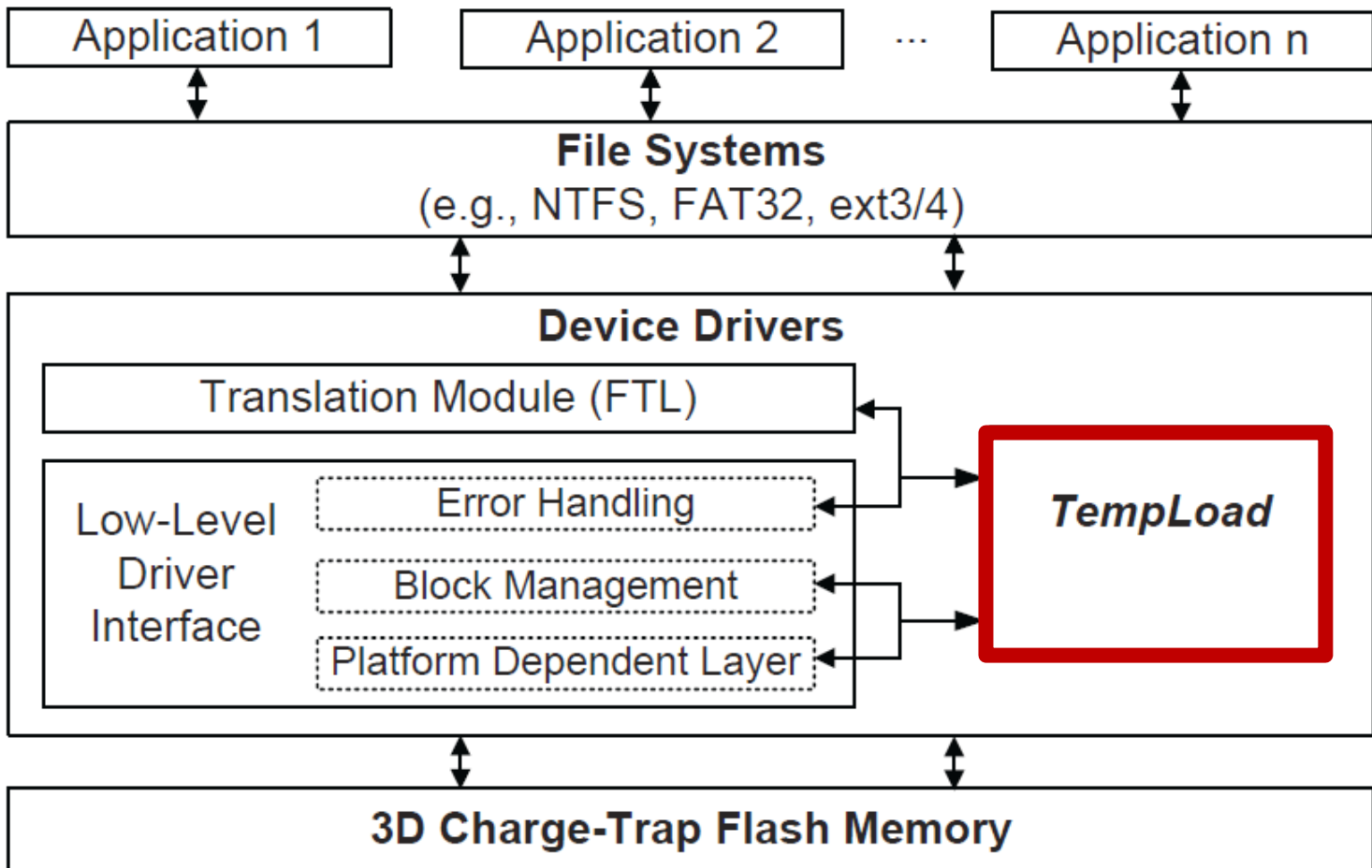
[1] M. K. Jeong, S. M. Joe, C. S. Seo, K. R. Han, E. Choi, S. K. Park, and J. H. Lee, "Analysis of random telegraph noise and low frequency noise properties in 3-D stacked nand flash memory with tube-type polysi channel structure," in *2012 Symposium on VLSI Technology (VLSIT)*, June 2012, pp. 55–56.

[2] M. Toledano-Luque, R. Degraeve, P. J. Roussel, V. S. Luong, B. H. Tang, J. G. Lisoni, C. L. Tan, A. Arreghini, G. V. den bosch, G. Groeseneken, and J. V. Houdt, "Statistical spectroscopy of switching traps in deeply scaled vertical poly-Si channel for 3D memories," in *2013 IEEE International Electron Devices Meeting (IEDM)*, Dec 2013, pp. 21.3.1–21.3.4.

TempLoad: A Temperature-Aware Data Allocation Strategy for 3D Flash Memory

- TempLoad allocates physical space based on the temperature status.
- TempLoad does not require prior knowledge of temperature.
- TempLoad does not need to maintain the temperature status of each physical block.
- **Objective:** reduce peak temperature and enhance data integrity

System Architecture of TempLoad



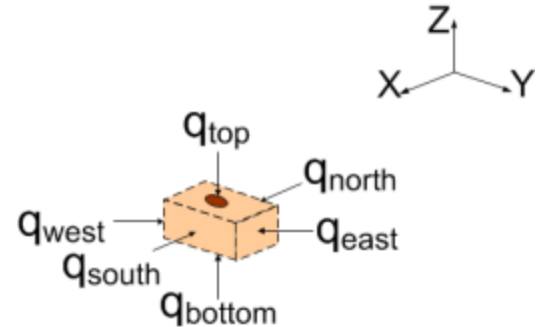
Reliability and Thermal Models

- We use lifetime reliability (LR) or instantaneous mean time to failure (MTTF) to model the impact on flash temperature [3].

$$LR \propto \left\{ \left[\ln\left(\frac{a}{1 + 2e^{b/kT}}\right) - \ln\left(\frac{a}{1 + 2e^{b/kT}} - c\right) \right] \times \frac{T}{e^{-d/kT}} \right\}^{\frac{1}{\delta}}$$

- The temperature differences can be modelled as [4-5]

$$C \frac{dT}{dt} = R^{-1} T(t) - pU(t)$$



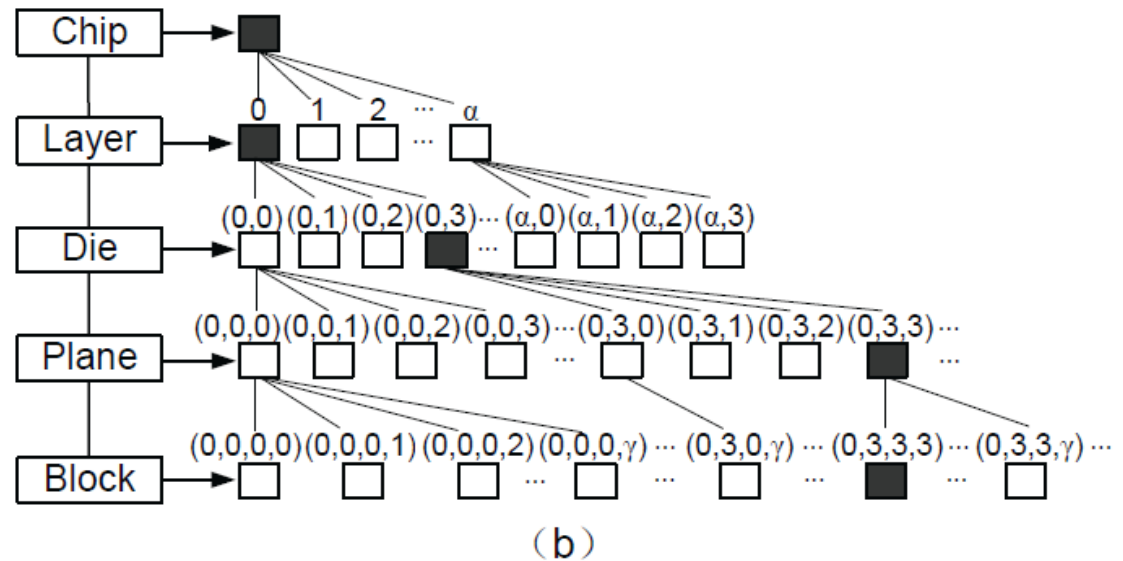
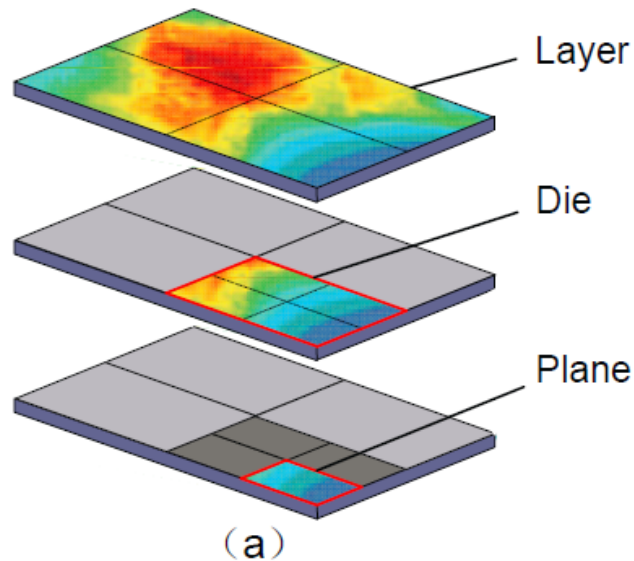
[3] J. Srinivasan, S. V. Adve, P. Bose, and J. A. Rivers, "Lifetime reliability: toward an architectural solution," *IEEE Micro*, vol. 25, no. 3, pp. 70–80, May 2005.

[4] A. Sridhar, A. Vincenzi, M. Ruggiero, T. Brunschweiler, and D. Atienza, "3D-ICE: Fast compact transient thermal modeling for 3D ICs with inter-tier liquid cooling," in *IEEE/ACM International Conference on Computer-Aided Design (ICCAD '10)*, 2010, pp. 463–470.

[5] A. Fourmigue, G. Beltrame, and G. Nicolescu, "Efficient transient thermal simulation of 3D ICs with liquid-cooling and through silicon vias," in *Design, Automation and Test in Europe Conference and Exhibition (DATE '14)*, 2014, pp. 1–6.

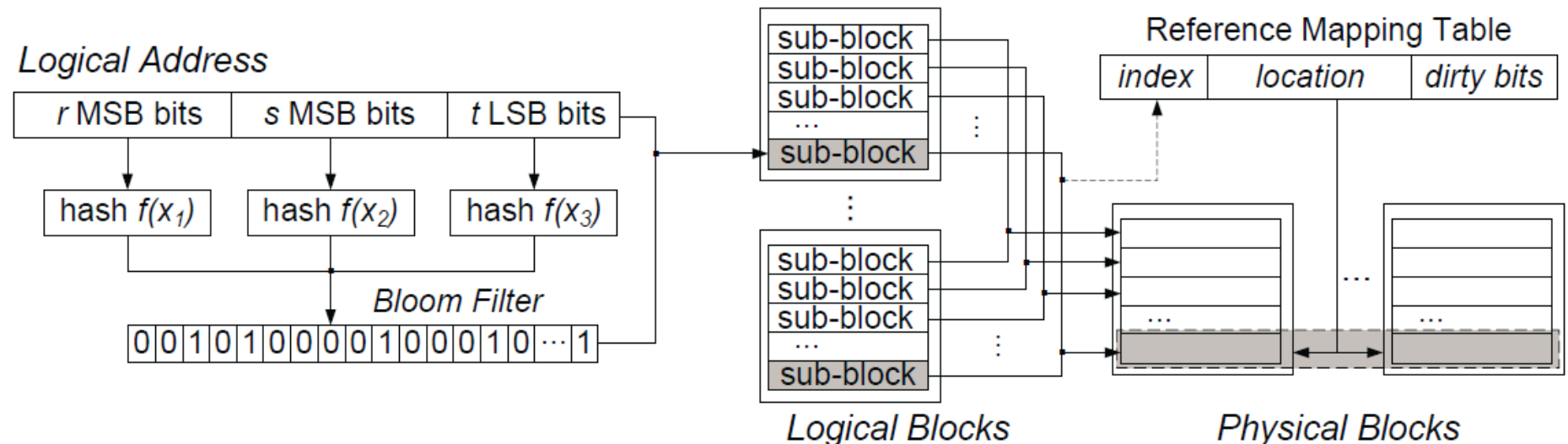
Block Selection with Temperature Mining

- Using this multiple-level quad tree, TempLoad only needs to compare the temperatures of a fixed number of sampling points at each stage.



Data Allocation with Set-Associative Block

- Solve the large-capacity block issue in 3D flash.
- Reduce the unnecessary garbage collection.



Evaluation

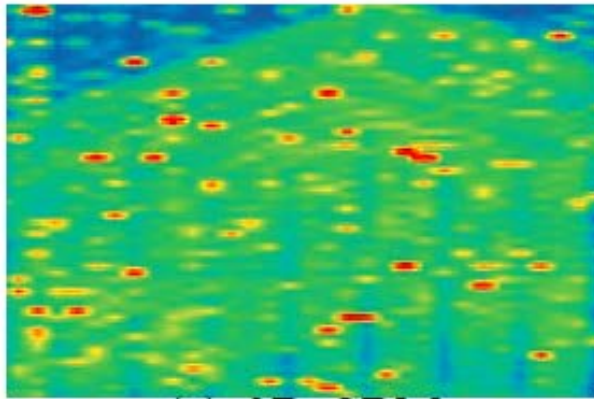
- TempLoad was implemented on the built-in NAND flash memory driver in Linux kernel 4.0.2.
- 3D-CBM [6] is selected as the baseline scheme.
- A 256Gb NAND flash memory is configured based on the specifications of a 3D flash memory test chip PF29F32B2ALCMG2 from Intel.

CHARACTERISTICS OF TRACES.

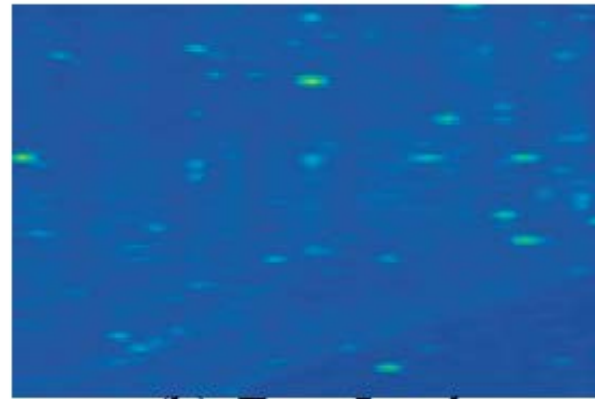
Trace	# of write operations	# of read operations	% of write	% of read
Financial	4,099,354	1,235,633	76.84	23.16
Webserver	1,260	4,260,449	0.03	99.97
onlineGames	653,133	817,887	44.40	55.60
KV-store	2,087,310	247,489	89.40	10.60

[6] Y. Wang, Z. Shao, H. Chan, L. Bathen, and N. Dutt, "A reliability enhanced address mapping strategy for three-dimensional (3-D) NAND flash memory," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 22, no. 11, pp. 2402–2410, Nov 2014.

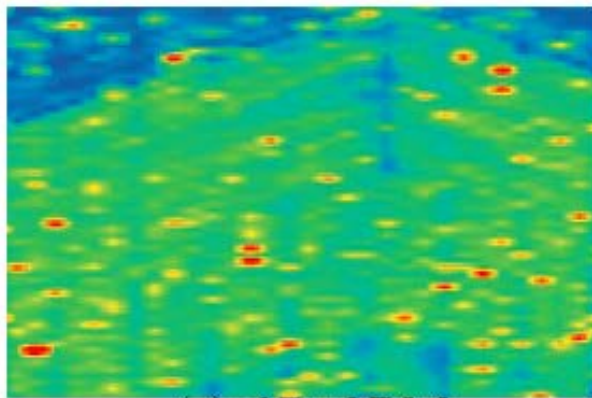
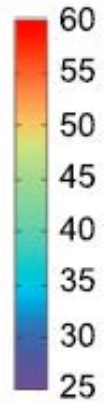
Peak and Average Temperature



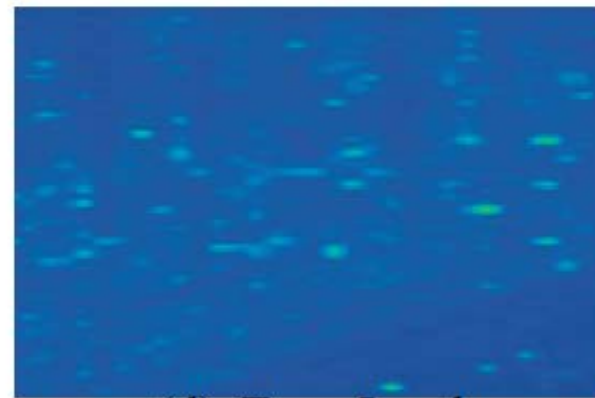
(a) 3D-CBM



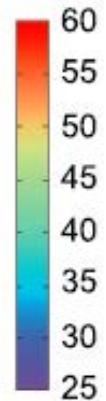
(b) TempLoad



(c) 3D-CBM

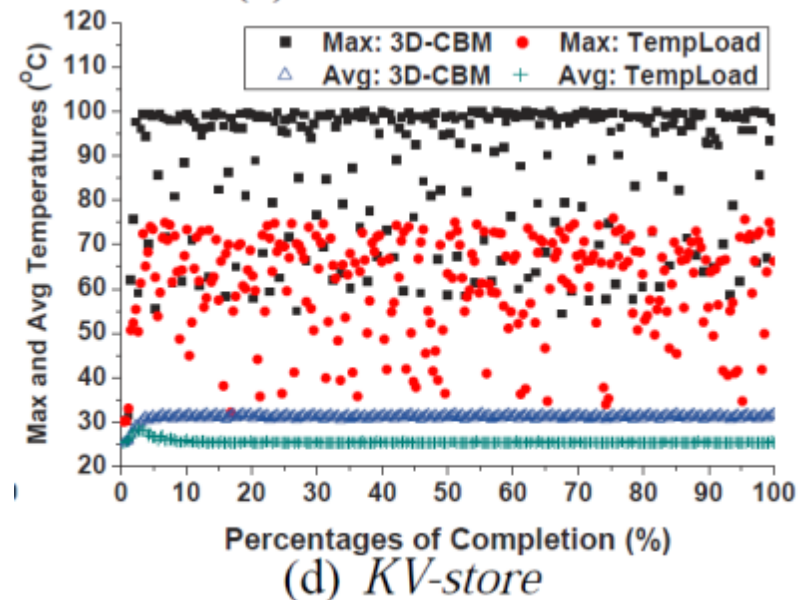
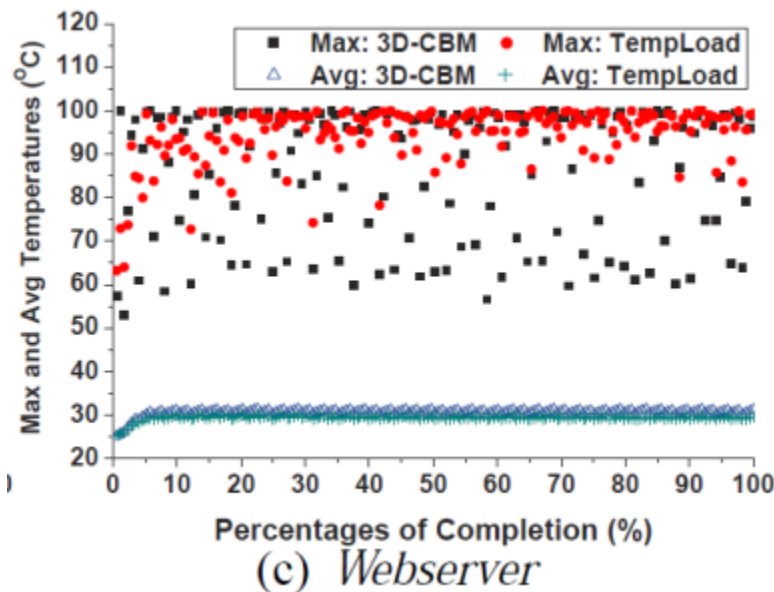
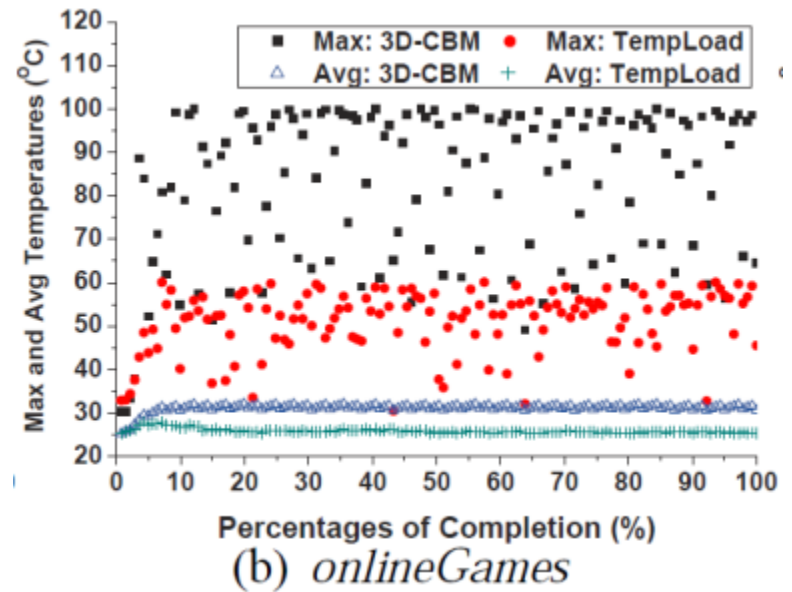
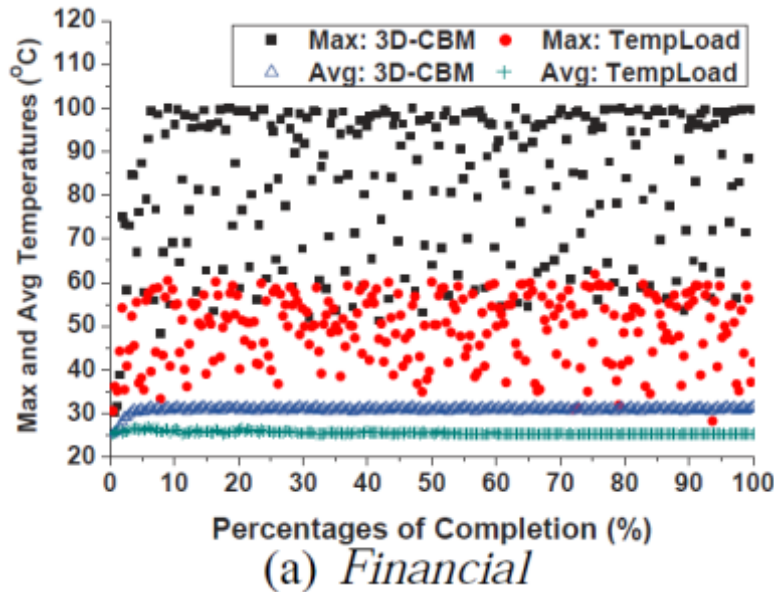


(d) TempLoad

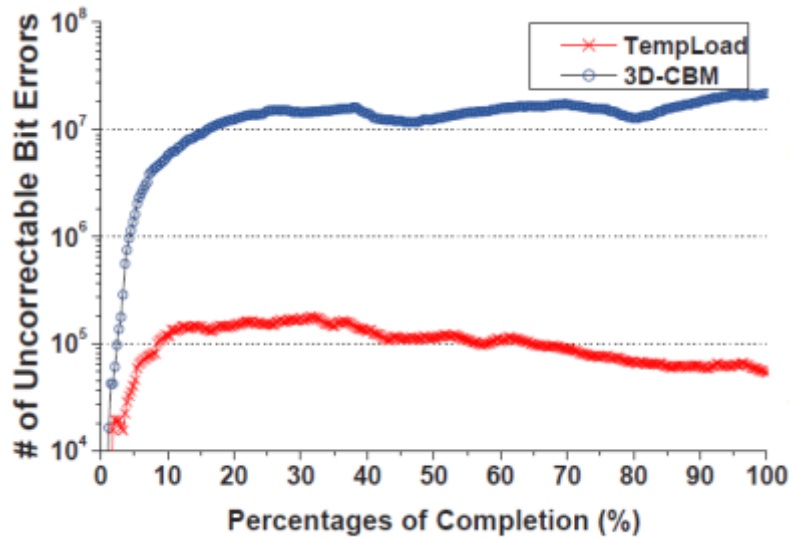


- The thermal image for standard benchmarks *Financial* (Figures (a) and (b)) and *onlineGames* (Figures (c) and (d)).

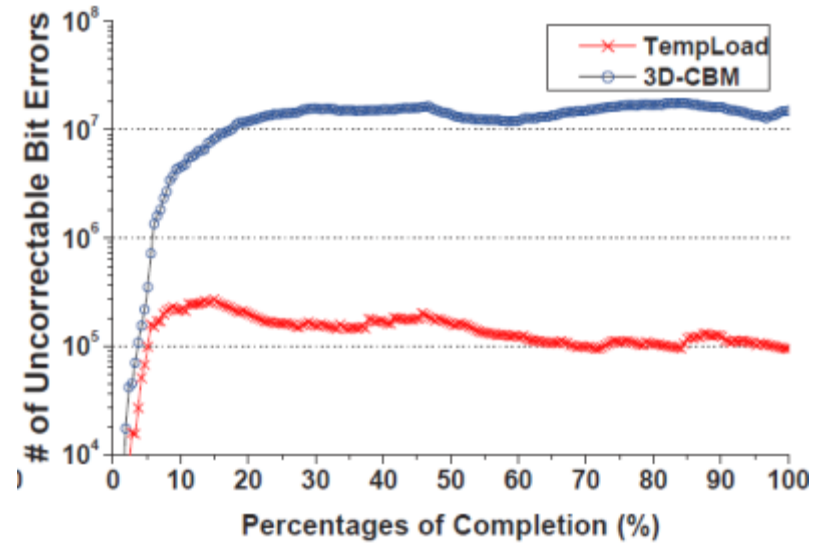
Peak and Average Temperature



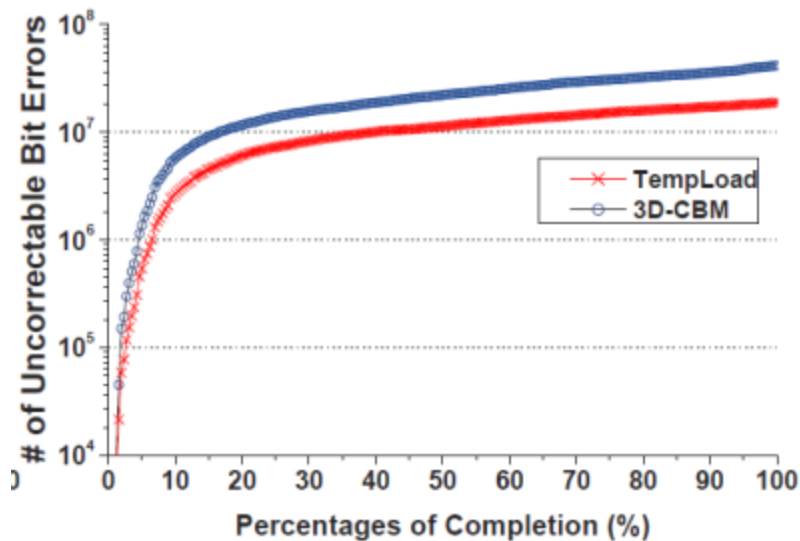
Data Integrity



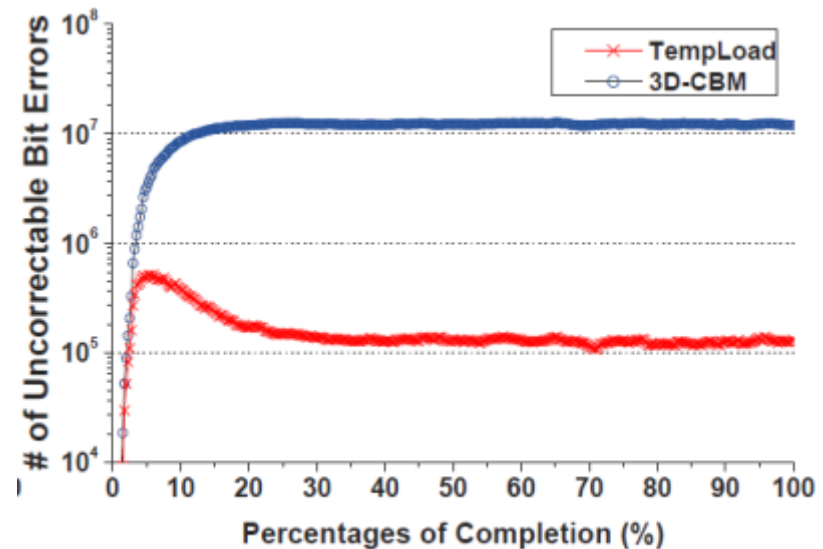
(a) *Financial*



(b) *onlineGames*



(c) *Webserver*



(d) *KV-store*

Response Time & Block Erase Counts

■ Response time

	3D-CBM (s)	TempLoad (s)
Financial	1.2469	0.9601
Webserver	1.0632	0.9620
onlineGames	2.5931	1.7078
KV-Store	3.3024	3.2281

■ Block erase counts

	3D-CBM	TempLoad
Financial	1,944,522	1,470,692
Webserver	1,096,940	955,315
onlineGames	1,174,240	787,942
KV-Store	2,284,932	2,172,934

Conclusion

- We present TempLoad, the first system-level temperature-aware data allocation strategy for 3D charge-trap flash memory.
- TempLoad allocates critical data to relatively low temperature physical blocks and improves the physical space utilization ratio.
- Experimental results show that TempLoad can significantly reduce the peak temperature and effectively enhance the data integrity

Thank you!