

異種の不揮発性メモリで構成される半導体ストレージシステムに関する研究

Research on Semiconductor Storage System with Heterogeneously Integrated Non-volatile Memories

情報セキュリティ科学専攻 松井 千尋

Chihiro MATSUI, Information Security Sciences

1. 論文の概要

新たな半導体不揮発性メモリであるストレージクラスメモリ (storage class memory, SCM) は次世代のコンピュータアーキテクチャを大きく変化させる。本論文では異なる種類の半導体不揮発性メモリを用いた、二種類のヘテロジニアスストレージおよびそのデータ管理アルゴリズムを提案する。頻繁にアクセスされるデータをメモリタイプ SCM あるいはストレージタイプ SCM に保存することで、NAND フラッシュメモリを用いたストレージを高速化することを実現した。さらに、ストレージアプリケーションによって適切な半導体不揮発性メモリの組み合わせが異なることを明らかにした。またエラー訂正符号を用い、半導体不揮発性メモリの高信頼化を実現した。図1に本論文の構成を示す。

2. 半導体不揮発性メモリの特性

図2にメモリおよびストレージ階層を示す。階層の上から下へアクセス性能は低下するが、メモリチップ当たりの容量は増加する。新たな半導体不揮発性メモリである SCM は、DRAM より大容量であり NAND フラッシュメモリより高速なアクセス性能を有する。磁気抵抗型メモリ (MRAM)、抵抗変化型メモリ (ReRAM)、相変化メモリ (PRAM) が主な SCM である。それぞれ材料や書き換え原理が異なるため、アクセス速度、容量などの特性が異なる。本論文では DRAM に近い高速なアクセス性能を持つ MRAM をメモリタイプ SCM (memory-type SCM, M-SCM) と呼び、NAND フラッシュと同等に大容量な ReRAM および PRAM をストレージタイプ SCM (storage-type SCM, S-SCM) と呼ぶ[1]。

NAND フラッシュメモリセルに 2 bit 以上を割り当てる多値化技術は、NAND フラッシュメモリの大容量化を推し進めてきた。Multi-level cell (MLC, 2 bit/cell), Triple-level cell (TLC, 3 bit/cell) NAND フラッシュメモリはメモリセルにそれぞれ 2 bit, 3 bit を割り当てる。NAND

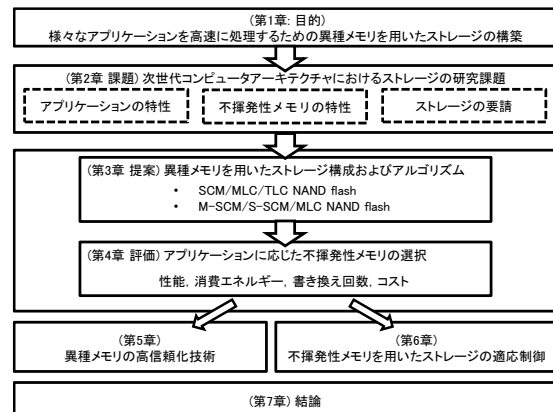


図1 本論文の構成

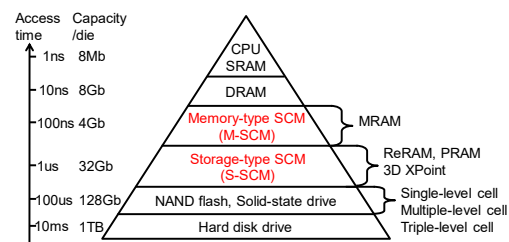


図2 メモリおよびストレージ階層

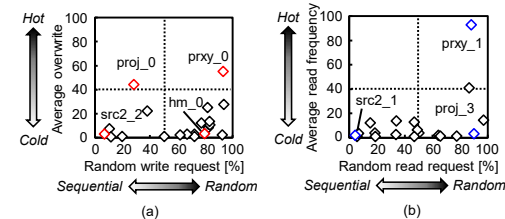


図3 ストレージアプリケーションの特性分類。(a)書き込みの多いアプリケーション, (b)読み出しの多いアプリケーション[3]

フラッシュメモリの多値化技術により大容量化する一方、しきい値分布が狭くなることでそのアクセス性能および信頼性は低下するという利害得失がある。

3. アプリケーション特性に適したストレージの必要性

表 I
半導体不揮発性メモリの特性[4]

Memory device	M-SCM (MRAM)	S-SCM (ReRAM, PRAM)	MLC flash	TLC flash
Read latency	0.1 us/sector	10 us/sector	52 us/page (U) 36 us/page (L)	80 us/page (U) 100 us/page (M) 80 us/page (L)
Write latency	0.1 us/sector	10 us/sector	2000 us/page (U) 370 us/page (L)	4400 us/page (U) 1500 us/page (M) 640 us/page (L)
Erase latency		Not required	3300 us/block	3200 us/block
I/O frequency		1066 MHz		400 MHz
V _{DD} (Core, I/O)		1.8V, 1.2V		3.3V, 1.8V
Access unit		Sector (512 byte)		Page (16 kbyte)
Acceptable endurance	10 ¹²	10 ⁸	10 ⁴	10 ³
Bit cost	10	4	1	2/3

U: Upper page, M: Middle page, L: Lower page of NAND flash

ストレージアプリケーションもまたその種類によってさまざまな特性を持つ。図 3 はプロキシデータベースサーバ、プロジェクト用ディレクトリなどから取得した 1 週間の読み出し・書き込みリクエストを含むストレージアプリケーション[2]を、書き込み・読み出し量の多寡 (write/read-intensive), 平均アクセス頻度 (hot/cold), 平均データサイズ (random/sequential)を用いて分類した[3]。このように複雑な特性を持つストレージアプリケーションに対し、従来研究では SCM および NAND フラッシュメモリを用いたハイブリッドストレージを提案した[3]。しかし、より複雑化した将来のストレージアプリケーションに対応するため、異種の半導体不揮発性メモリを用いたヘテロジニアスストレージを構築する必要がある。ヘテロジニアスストレージの記憶には、表 I に示す特性を有する M-SCM, S-SCM, MLC および TLC NAND フラッシュメモリを用いる。

4. 異種の不揮発性メモリを用いたヘテロジニアスストレージ

図 4 にヘテロジニアスストレージを示す。ストレージアプリケーションに含まれる頻りにアクセスされるデータを高速な M-SCM あるいは S-SCM に、アクセス頻度の低いデータを低速・大容量な MLC あるいは TLC NAND フラッシュメモリに保存する。図 5 に二種類のデータ管理アルゴリズムを提案する。

第一の SCM, MLC および TLC NAND フラッシュメモリを用いたヘテロジニアスストレージは、MLC および TLC NAND フラッシュメモリの低い書き換え性能を隠ぺいするため、書き込みに適したデータ管理アルゴリズムを提案する[5]。図 5(a)に示すように書き込みデータの書き換え頻度に着目し、データ書き換え頻度の低いコールドデータを SCM から MLC NAND フラッシュメモリに移動する。さらに MLC NAND フラッシュメモリのガベージコレクションのタイミングで、書き換え頻度の低い

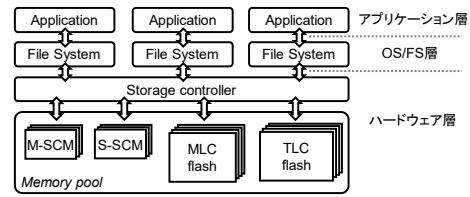


図 4 異種の半導体不揮発性メモリを用いたヘテロジニアスストレージの構成

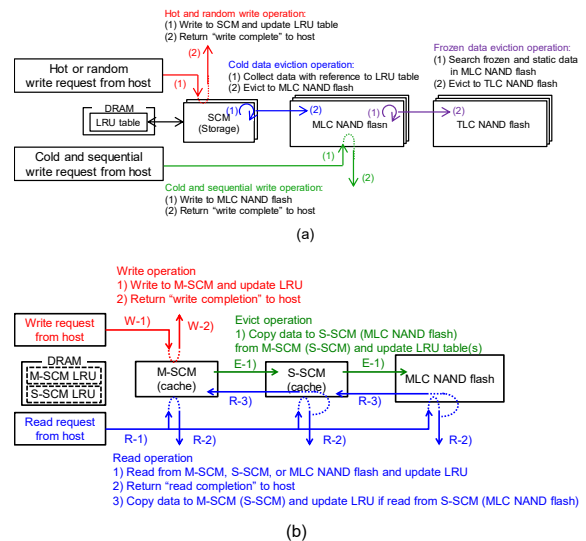


図 5 (a) SCM, MLC および TLC NAND フラッシュメモリを用いたヘテロジニアスストレージに適用する Cold and frozen data eviction (CFDE) [5]. (b) M-SCM, S-SCM および MLC NAND フラッシュメモリを用いたヘテロジニアスストレージに適用する 2 non-volatile memory write-back (2NV-WB) キャッシュ[6]

データを MLC NAND フラッシュメモリのから TLC NAND フラッシュメモリに移動する。これを Cold and frozen data eviction (CFDE) アルゴリズムと呼ぶ。

第二の M-SCM, S-SCM および MLC NAND フラッシュメモリを用いたヘテロジニアスストレージは、高速な M-SCM および S-SCM から読み出し・書き込みを行うよう、読み出し・書き込みに適したデータ管理アルゴリズムを提案する[6]。そのために二種類の SCM を用いるヘテロジニアスストレージ向けに、2 non-volatile memory write-back (2NV-WB) キャッシュアルゴリズム[図 5(b)]を提案した。下位の不揮発性メモリ (S-SCM あるいは MLC NAND フラッシュメモリ) へ保存されたデータに読み出し・書き込みリクエストがあると、上位の不揮発性メモリ (M-SCM あるいは S-SCM) へデータをコピーする。そのため図 5(a) に示した CFDE と異なり、図 5(b) に示す 2NV-WB は不揮発性

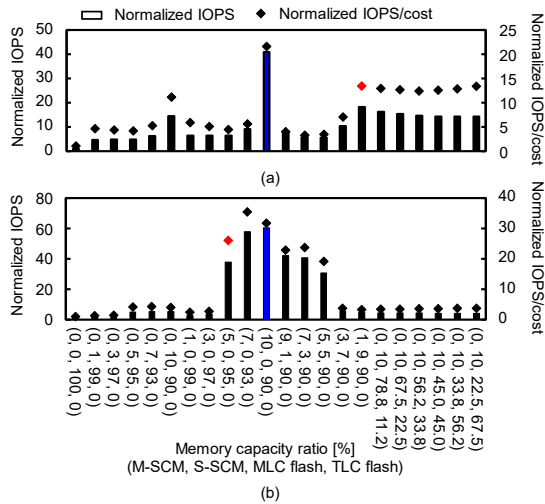


図 6 ストレージアプリケーションに最適なメモリ構成の評価。(a) prxy_0, (b) prxy_1 アプリケーション[4, 6]

表 II

ストレージアプリケーション特性に適切なヘテロジニアスストレージの構成

Application characteristic	Best non-volatile memory combination
Write-hot-random (prxy_0)	(1%, 9%, 90%, 0%)
Read-hot-random (prxy_1)	(5%, 0%, 95%, 0%)
Write-hot-sequential (proj_0)	(0%, 10%, 22.5%, 67.5%)
Write-cold-random (hm_0)	(0%, 10%, 78.8%, 11.2%)
Cold-sequential (src2_2, src2_1)	(1%, 0%, 99%, 0%)
Read-cold-random (proj_3)	or (0%, 0%, 100%, 0%)

メモリ間でデータを二重に持つ場合がある。揮発性メモリを用いるキャッシュマネジメントと異なり、不揮発である SCM を用いた提案の CFDE および 2NV-WB アルゴリズムは下位のメモリへ定期的なデータフラッシュが不要である。

5. アプリケーションに適したヘテロジニアスストレージの構成

ヘテロジニアスストレージに用いる半導体不揮発性メモリの容量比を変化させ、ストレージ性能(Input/Output per second, IOPS)の評価し、ストレージアプリケーションに最適なヘテロジニアスストレージの構成を明らかにする。SystemC を用いたストレージエミュレータに図 5 に示したデータ管理アルゴリズムを実装した。表 I に示す半導体不揮発性メモリの特性および図 3 に示すストレージアプリケーションのうち代表的な 7 種をエミュレータの入力とする。

M-SCM 容量および S-SCM 容量の和を、ヘテロジニアスストレージの総容量の 10%以内と仮定する。図 6 に prxy_0 および prxy_1 アプリケーションの IOPS 性能を示す。これらのス

トレージアプリケーションは、図 3 の特性分類によるとそれぞれ書き込み、読み出しが多く、平均的なアクセス頻度が高く平均データサイズが小さい。全体のストレージ容量に対する M-SCM あるいは S-SCM の容量比が同じでも、M-SCM と S-SCM のアクセス性能の差により、高速な M-SCM 容量が多いほど、ストレージ性能が向上することを明らかにした。また総ストレージコストを式(1)で定義すると、

$$\begin{aligned} \text{Total storage cost} \\ = \Sigma(\text{memory capacity ratio} \\ \times \text{bit cost ratio}) \end{aligned} \quad (1)$$

表 I のビットコスト比から、M-SCM あるいは S-SCM 容量が多いほど総ストレージコストが上昇する。総ストレージコストを考慮し IOPS/cost という指標を導入すると、図 6(a)に示す prxy_0 アプリケーションの場合、M-SCM は小容量用い、S-SCM は大容量用いることが適切であることを明らかにした。一方で図 6(b)の prxy_1 アプリケーションでは、高速な M-SCM をできるだけ大容量用いると性能を向上できることを明らかにした。

図 3 に示した他の 5 アプリケーションについても同様に、アプリケーションに特性に適切なヘテロジニアスストレージの構成を表 II に示す。M-SCM が必要なアプリケーションは、ホット・ランダムなアプリケーション(prxy_0, prxy_1)である。書き込みが多く、ホットあるいはランダムなアプリケーション(proj_0, hm_0)には S-SCM, MLC および TLC NAND フラッシュメモリを用いたヘテロジニアスストレージが適している。しかしコールドなアプリケーション(src2_1, src2_2, proj_3)については、SCM による性能向上が限定的であるため SCM を用いる必要は無い。またストレージ性能の高い不揮発性メモリの組み合わせほど、M-SCM あるいは S-SCM へ多くアクセスされるため、これに反比例してストレージの消費エネルギーは低いことを明らかにした[5, 6]。

6. エラー訂正符号を用いたストレージの高信頼化

半導体不揮発性メモリは書き換えやデータ保持によってメモリセルにエラーが発生するため、エラー訂正符号(error-correcting code, ECC)を用いて高信頼化を実現する。ランダムエラーを訂正する Bose-Chaudhuri-Hocquenghem (BCH)符号を用い、M-SCM および MLC NAND フラッシュメモリを用いたストレージを高信頼化し、そのときの性能を評価した。図 7(a)に示すように、読み出しの多いアプリケーション(prxy_1)は M-SCM の訂正可能ビットを増やすほど性能が低下することを明ら

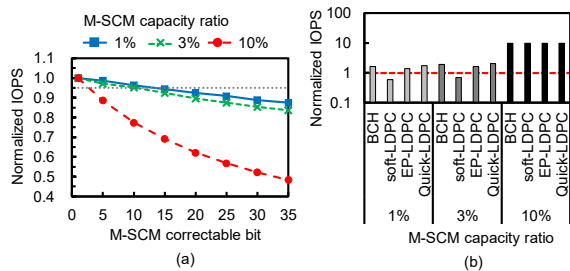


図 7 エラー訂正符号を用いた (a) SCM の高信頼化, (b) MLC NAND フラッシュメモリの高信頼化[8]

かにした。また、M-SCM 容量が多いほど M-SCM から読み出されるデータ量が増えるため、ECC により性能が低下する。そのため、ウェアレベリングによる書き換え回数の均一化、データリフレッシュによるデータ保持時間の均一化などを用いることで、SCM の必要な ECC 訂正能力を低く抑えることが必要である。

一方で NAND フラッシュメモリは多値化およびスケールリングにより多くのエラーが発生するため、より強力な ECC である low-density parity-check (LDPC) 符号を用いることが必要と考える。Error-Prediction (EP-) LDPC 符号および Quick-LDPC 符号[7]を用い、LDPC 符号に必要な読み出し回数を減らしつつ信頼性を担保する。図 7(b)に MLC NAND フラッシュに符号化率 9/10 の BCH および LDPC 符号を適用したときのストレージ性能を示す。SCM 容量を増やすほど MLC NAND フラッシュメモリから読み出されるデータ量が減少するため、図 7(b)に示すように MLC NAND フラッシュメモリには BCH と比較して訂正能力の高い LDPC 符号を適用することができることを明らかにした[8]。

7. まとめと今後の展望

本研究では、新たな半導体不揮発性メモリである SCM および NAND フラッシュメモリの特性を考慮し、異種の半導体不揮発性メモリを用いたヘテロジニアスストレージおよびそのデータ管理アルゴリズムを提案した。ストレージアプリケーション特性によって適切なメモリの構成は異なり、特に頻繁にアクセスがあり平均データサイズの大きいアプリケーションには高速な M-SCM が必要であることを明らかにした。一方で低速だが大容量の TLC NAND フラッシュメモリが必要なストレージアプリケーションが存在することも明らかにした。また ECC を用い、半導体不揮発性メモリの高信頼化技術を実現した。SCM の特性およびストレージアプリケーションにより、ECC を適切な ECC の強度があることを明らかにした。

3次元積層された NAND フラッシュメモリの特性を考慮したデータマネジメントアルゴリズムや、三種以上の半導体不揮発性メモリを用いたヘテロジニアスストレージの高信頼化技術を検討することが今後の課題である。

関連する発表文献

学術雑誌 筆頭著書 4 件, 共著書 4 件
 国際会議 筆頭著書 5 件, 共著書 9 件
 国内会議 筆頭著書 4 件, 共著書 4 件
 その他 筆頭著書 2 件, 共著書 1 件

参考文献

- [1] IBM Almaden Research Center, "Storage class memory: Towards a disruptively low-cost solid-state non-volatile memory," http://researcher.watson.ibm.com/researcher/files/us-gwburr/Almaden_SCM_overview_Jan_2013.pdf, Jan. 2013.
- [2] MSR Cambridge Traces, <http://iotta.snia.org/traces/388>.
- [3] S. Okamoto, C. Sun, S. Hachiya, T. Yamada, Y. Saito, T. O. Iwasaki, and K. Takeuchi, *IEEE International Memory Workshop*, May 2015, Monterey, pp. 157-160.
- [4] C. Matsui, T. Yamada, Y. Sugiyama, Y. Yamaga, and K. Takeuchi, "Optimal memory configuration analysis in tri-hybrid solid-state drives with storage class memory and multi-level cell/triple-level cell NAND flash memory," *Japanese Journal of Applied Physics (JJAP)*, vol. 56, no. 4S, pp. 04CE02-1 - 04CE02-9, Apr. 2017.
- [5] C. Matsui and K. Takeuchi, "8.9-times performance improvement by tri-hybrid storage system with SCM and MLC/TLC NAND flash memory," *International Conference on Solid State Devices and Materials (SSDM)*, Tsukuba, Sep. 2016, pp. 105-106.
- [6] C. Matsui and K. Takeuchi, "22% higher performance, 2x SCM write endurance heterogeneous storage with dual SCM and NAND flash memory," *European Solid-State Device Research Conference (ESSDERC)*, Leuven, Sep. 2017, pp. 6-9.
- [7] T. Tokutomi, M. Doi, S. Hachiya, A. Kobayashi, S. Tanakamaru, and K. Takeuchi, "Enterprise-grade 6x fast read and 5x highly reliable SSD with TLC NAND-flash memory for big-data storage," *IEEE International Solid-State Circuits Conference*, San Francisco, Feb. 2015, pp. 140-141.
- [8] C. Matsui, R. Kinoshita, and K. Takeuchi, "Analysis on applicable ECC strength of SCM and NAND flash in hybrid storage," *Japanese Journal of Applied Physics (JJAP)*, to be published in vol. 57, no. 4S, Apr. 2018.