

# データ断片化防止アルゴリズムを用いた ハイブリッド SCM/MLC NAND フラッシュ SSD

岡本 峻<sup>†</sup> 藤井 裕大<sup>†</sup> 宮地 幸祐<sup>‡</sup> 上口 光<sup>†</sup> 樋口 和英<sup>†</sup> 孫 超<sup>†</sup> 竹内 健<sup>†</sup>

データ断片化防止アルゴリズムを用いた SCM/MLC NAND フラッシュハイブリッド SSD を提案し、その SSD の速度、電力、書換え回数を評価した。使用頻度の高く、且つ、断片化しているデータは SCM に、断片化していないデータは NAND フラッシュに格納することで、部分書換えを抑制し消去回数を低減した。さらに NAND フラッシュ内のデータに部分的な書換えが発生した場合には SCM に書き戻す。提案手法により、NAND フラッシュ内のデータの断片化を防ぎ、SSD 書込み性能が 11 倍向上した。

## 1. はじめに

ビッグデータを活用する上で、ストレージの高速アクセスを可能とする SSD (Solid State Drive) が注目されている。SSD は、現在広く利用されている HDD (Hard Disk Drive) よりも高速、低電力であるが、記憶素子に用いられる NAND フラッシュメモリの微細化限界や消去と書込み単位の不一致による非効率な電力・容量消費が問題となっている。これらの問題を解決するために、抵抗変化型の不揮発メモリ、SCM (Storage Class Memory) に注目が集まっている。SCM は、DRAM のように高速で不揮発性メモリの性質を持つが、現状では NAND フラッシュと比較して容量コストは高い。そこで、SCM と NAND フラッシュをハイブリッドで搭載した SSD を開発し、限られた容量の SCM を有効に活用するアルゴリズムを提案した。

## 2. ハイブリッド SCM/MLC NAND SSD

SCM は、NAND フラッシュより低消費電力で、読出し/書込み時間も遥かに高速 (数十 ns/bit) な新型不揮発性メモリである。NAND フラッシュメモリは、その構造的制約から、書き込みと消去の単位が異なるため、上書きができない。そのため、部分書換えが起こると別のページにそのデータをコピーし、元のページを無効化する必要がある。従って、部分書換えが頻発すると、書込みデータが断片化するだけでなく、無効ページを再利用するためのブロック回収により、急激に書込み性能が低下する。一方、SCM は書込み単位が NAND フラッシュメモリより小さい (~数十 bit) ため、上書きが可能である。SCM には、ReRAM (抵抗変化型メモリ)、PRAM (相変化メモリ)、MRAM (磁気抵抗メモリ) などがある。

本論文では、SCM/MLC NAND フラッシュハイブリッド SSD を提案する。ハイブリッド SSD は、SSD コントローラの下に NAND フラッシュチップと SCM チップを実装する (図 1)。SSD コントローラでは論理・物理アドレス変換、ウェアレベリング、ガベージコレクション等を行い、DRAM に各種テーブルを展開する。

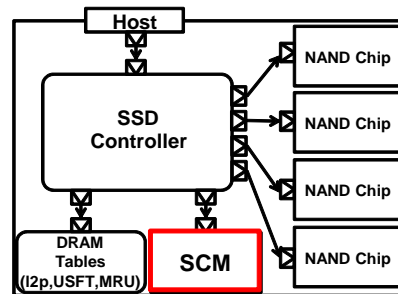


図 1 3次元ハイブリッド SCM/NAND フラッシュ SSD の構造

## 3. データ断片化防止アルゴリズムの評価

NAND フラッシュは、書き込み単位であるページサイズ (8 Kbyte - 16 KByte) とオペレーティング・システム (OS) からの書込み命令の単位であるセクタサイズが一致していない。これにより、ページサイズ以下のデータが書き込まれるとページの断片化が生じ、書換え回数の増加の原因となる。その結果、ブロック回収が生じ、消費電力の増大や書込み性能の低下に繋がる。

そこで、データ断片化防止アルゴリズムを提案した。断片化の評価には、使用セクタテーブルを用いる。使用セクタテーブルは、有効なセクタにフラグを立て、有効フラグ数から算出したページの使用率により、断片化を判別する。算出したページの使用率が閾値を下回っていれば、書込みデータは断片化していると判断し、SCM に書込む。そうでない場合には、NAND フラッシュに優先的に書込む。図 2 に使用セクタテーブルの更新手順を示す。Case 1 のように新しい領域に書く場合、ページ使用率の情報を持たないので、書込みデータからページ使用率を計算し、使用セクタテーブルに格納する。Case 2 のようにデータの追記や上書きが発生する場合、保持しているデータのページ使用率と書込みデータのそれとをマージして使用セクタテーブルを更新する。

Algorithm flow	例. case 1	例. case 2
ページ使用状況		
ホストからの書き込み		
最新のページ使用状況		
ページ使用率 $R$	$R = 7/8 = 0.875$	$R = 4/8 = 0.5$
閾値 $R_{TH}$ : (例. 0.7)	$R \geq R_{TH}$ 断片化していない	$R < R_{TH}$ 断片化している
データ書き込み先	Write to MLC NAND	Write to ReRAM
	□ Empty "0" ■ Used "1"	

図2 使用セクタテーブルの更新

更に、SCM を用いて、データアクセスの時間的局所性を有効に活用するため、MRU テーブルを実装した。MRU テーブルは、新規アクセスの場合には、FIFO の要領で書き込み、既に書き込みデータのアドレスが存在する場合には、そのアドレスを先頭に移動し、並び替える。先の断片化防止アルゴリズムにおいて、データが断片化していないと判断された場合でも、MRU テーブルを参照し、書き込むアドレスが MRU テーブルに存在すれば、書き込みデータは書換え頻度が高いとみなし、SCM に格納する。

NAND フラッシュに一旦書込まれたデータでも、頻繁に断片化した小さい単位のアクセスがある場合には、断片化データとして、SCM に書き戻す。

#### 4. 結果

TLM (Transaction Level Modeling) ベースの SSD エミュレータを用いて、提案したハイブリッド SCM/MLC NAND フラッシュ SSD の書き込み性能、エネルギー消費量、書き込み/消去 (P/E) サイクルを評価した。評価には、金融サーバのアクセスパターン (financial1) を用い、比較のために、従来の MLC NAND フラッシュのみの SSD も評価した。図 3、図 4 に示す通り、MLC NAND のみの SSD と比較すると、提案ハイブリッド SSD は、書き込み性能が 11 倍向上し、書き込みエネルギーが 79% 減少した。3 次元 TSV 配線を用いれば、I/O エネルギーは 1/27 に減少するので、結果として、書き込みエネルギーは、93% 削減できる。さらに、図 5 に示すように、P/E サイクルの傾き平均は、6.9 倍減少するので、提案 SSD では、寿命が 6.9 倍長くなる。

#### 5. 結論

ハイブリッド SCM/MLC NAND フラッシュ SSD のアーキテクチャを提案した。提案した SSD データ管理アルゴリズムを用いることで MLC NAND フラッシュのデータ断片化を抑制し、SCM にページサイズに対して小さいデータと頻繁に書換えられるデータを格納することで、効率的に SCM を活用した。提案したハイブリッド SSD では、従来の MLC NAND フラッシュのみの SSD より、11 倍書き込み性能が増加し、6.9 倍寿命が延び、93% 書き込みエネルギーが削減された。

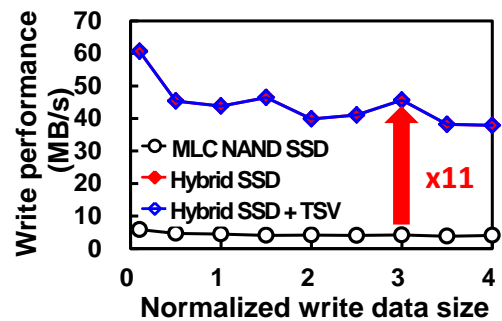


図3 書き込み性能

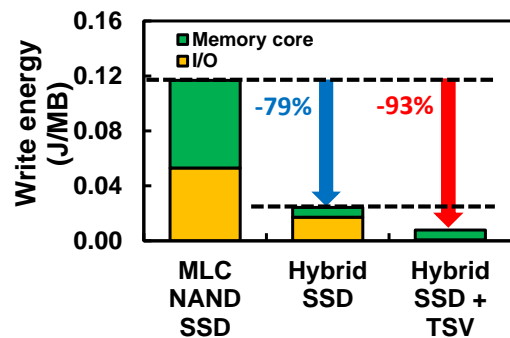


図4 書き込みエネルギー

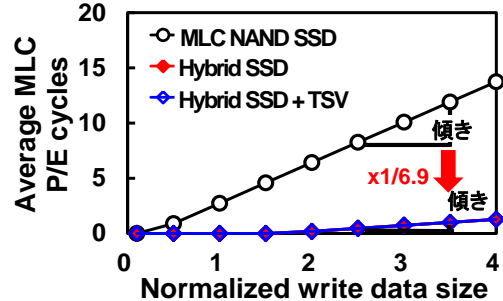


図5 Program/Erase サイクル

#### 謝辞

本研究は NEDO の助成により行われました。

#### 参考文献

- [1] H. Fujii et al, "x11 performance increase, x6.9 endurance enhancement, 93% energy reduction of 3D TSV-integrated hybrid ReRAM/MLC NAND SSDs by data fragmentation suppression," *IEEE Symp. on VLSI Circuit*, pp134-135, June 2012.
- [2] T. Hatanaka et al, "A 60% higher write speed, 4.2Gbps, 24-channel 3D-solid state drive (SSD) with NAND flash channel number detector and intelligent program-voltage booster," *IEEE Symp. on VLSI Circuits*, pp. 233-234, June 2010.