

# ガベージコレクション回数を削減し NAND フラッシュ SSD を高速化するミドルウェア

松井 千尋<sup>1</sup> 曾我 あゆみ<sup>1</sup> 孫 超<sup>1</sup> 竹内 健<sup>1</sup>

ファイルシステムと NAND フラッシュ SSD の Flash Translation Layer の間でアドレス変換を媒介するミドルウェア LBA scrambler を提案する。LBA scrambler は次にガベージコレクションを行うおとする NAND フラッシュのブロックに意図的にデータを書き込むようにアドレス変換を行う。提案手法によりガベージコレクション回数を削減し、書き込み性能が最大 2.2 倍向上した。

## 1. 研究の目的, 背景

高速, 低電力な NAND フラッシュ SSD (solid-state drives) は, NAND フラッシュメモリが読み込み・書き込みサイズ (ページ) と消去サイズ (ブロック) が異なること, 上書き前に消去が必要であることによって, 書き込み性能が制限される。そのため, Flash Translation Layer (FTL) が論理アドレスから物理アドレスへの割り当てや有効ページの管理等を行う。また, SSD 内の空ブロック数が少なくなると, ガベージコレクション (GC, garbage collection) が発動する。回収するブロック内の有効ページを別のブロックにコピーすることによって空のブロックを確保する (図 1)。ブロック内の有効ページ数が多い場合には GC に 100 ミリ秒を超える時間がかかり, SSD の書き込み性能を悪化させる。この問題を解決するため, 図 2 に示すように, 次に消去するブロックに積極的にデータを書き込むことで GC の回数を削減する手法を提案する。

## 2. LBA scrambler を配置した SSD のストレージシステム

新たなレイヤ logical block address (LBA) scrambler をファイルシステム (OS) と FTL との間の中ルウェアとして導入する (図 3)。LBA scrambler は FTL から推奨される次に消去するブロックのページアドレスを上書き推奨リストとし

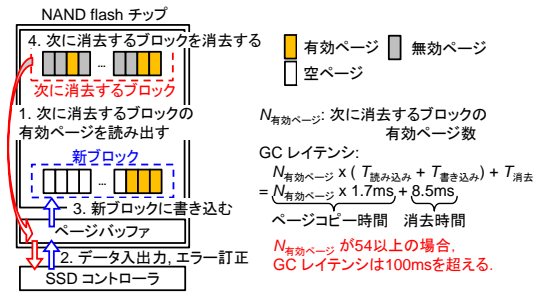


図 1 NAND フラッシュの GC 動作[1]

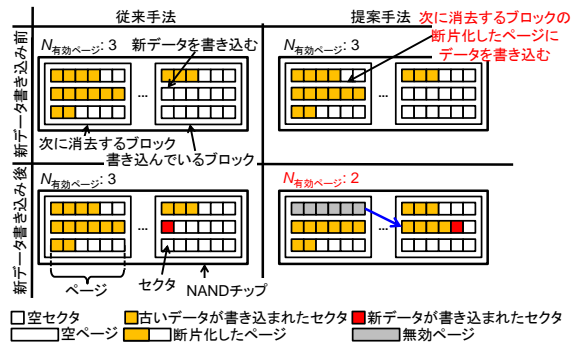


図 2 GC 回数を削減する提案手法[2]

て持つ。上書き推奨リストに基づいて OS から送られてきた LBA を SLBA (scrambled LBA) に変換し, FTL に書き込みリクエストを送る。FTL は SLBA を物理アドレス (physical address) に変換し, NAND に書き込みを行う。LBA scrambler は上書き推奨リストに加えて, LBA から SLBA への変換テーブル, 未使用 SLBA のテーブルを持つ。

<sup>1</sup> 中央大学理工学部

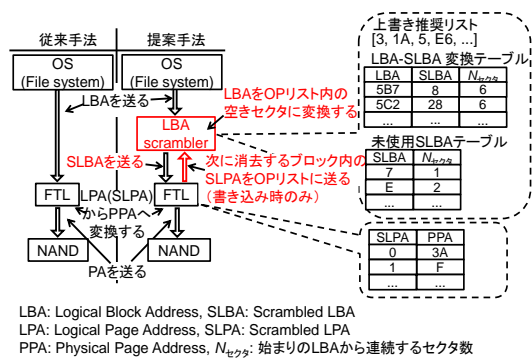


図3 LBA scambler を配置した SSD のストレージシステム[2]

表1 NAND フラッシュのパラメータ設定[2]

ワークロード	Financial1	tpcc-mysql	prxy_0	prxy_1	src1_0
データサイズ[GB]	0.53	0.65	0.84	13.75	60.62
SSD空き容量	20%, 40%, 60%, 80%				
NAND容量	データサイズ / (1 - SSD空き容量)				
ブロック内ページ数	256				
ブロック、ページサイズ	4MB, 16KB				

提案手法において FTL は SLBA を LBA とまったく同じに扱うため、他のシステムに変更は必要ない。

### 3. LBA scambler のアルゴリズム

FTL から与えられた上書き推奨リストに基づいて、LBA scambler は新しい書き込みリクエストを次の3つのルールに従い作成する。

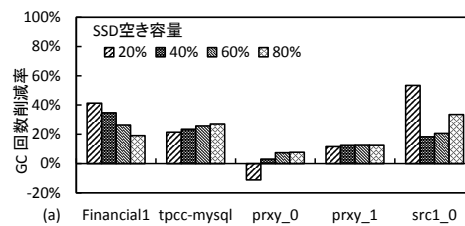
ルール(1) SSDの空き容量が十分な場合、空のページに書き込む。

ルール(2) 次にガベージコレクションを行うブロックに十分な空きページがある場合、そのブロックの断片化したページに書き込む。

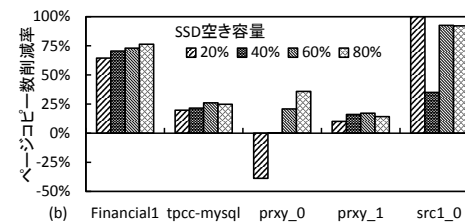
ルール(3) 次にガベージコレクションを行うブロックに十分な空きページがない場合、そのブロック以外の空ページあるいは断片化したページに書き込む。

### 4. LBA scambler を配置したシステムの評価

提案する LBA scambler を配置したエミュレータを設計し、SSDの書き込み性能を評価した。評価には表1に示す5種類のワークロード[3,4,5]を用い、従来手法との比較を行った。提案手法を用いたSSDは、GC回数およびGC中のページコピー数がそれぞれ削減した(図4(a), (b))。この結果図5に示すように書き込み性能が向上し、特に src1\_0 では



(a) Financial1 tpcc-mysql prxy\_0 prxy\_1 src1\_0



(b) Financial1 tpcc-mysql prxy\_0 prxy\_1 src1\_0

図4(a) GC回数削減率 (b) GC中のページコピー回数削減率

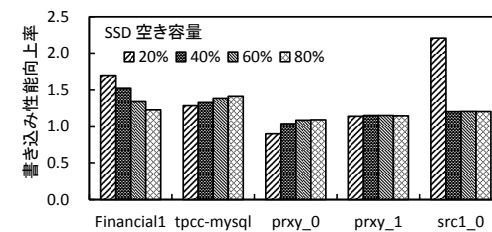


図5 SSD書き込み性能向上率[2]

2.2倍向上した。

## 5. 結論

ファイルシステムとNANDフラッシュSSDとの間に位置しアドレス変換を媒介するLBA scamblerを提案した。提案した手法によってSSDのGC回数およびGC中のページコピー回数が削減し、この結果書き込み性能が最大2.2倍増加した。

### 参考文献

- [1] K. Takeuchi et al., "A 56 nm CMOS 99mm<sup>2</sup> 8Gb multi-level NAND flash memory with 10Mbyte/sec program throughput", IEEE J. of Solid-State Circuits, vol. 42, no. 1, pp. 219-232, 2007.
- [2] A. Soga et al., "NAND Flash Aware Data Management System for High-Speed SSDs by Garbage Collection Overhead Suppression", IEEE International Memory Workshop, pp. 95-98, Taipei, Taiwan, May 2014.
- [3] <http://traces.cs.umass.edu/index.php/Storage/Storage>.
- [4] tpcc-mysql, <https://code.launchpad.net/~percona-dev/perconatools/tpcc-mysql>.
- [5] MSR Cambridge Traces, <http://iotta.snia.org/traces/388>.