

# SSDの発展を支える技術

Technologies Supporting Evolution of SSDs

土屋 憲司

■TSUCHIYA Kenji

SSD (ソリッドステートドライブ) は、可動部分がないためにHDD (ハードディスクドライブ) に比べて耐振動性や耐衝撃性に優れるという特長を持つ。この利点を生かし、SSDを搭載したノートPC (パソコン) が2006年頃から市場に登場するようになった。しかし、当時のSSDにはSLC (Single Level Cell) NANDフラッシュメモリが用いられており、SLCに比べて容量を2倍にできるMLC (Multi Level Cell) NANDフラッシュメモリは、信頼性などの技術的課題があるためにSSDには使用できないと信じられていた。

東芝は、ノートPCでのHDDの実際の使用状況を解析し、実ワークロードに適応したNANDキャッシュシステム、及び新NAND管理システムを開発することでSSDの信頼性に関わる書込み効率 (WAF) を低く抑え、MLC NANDフラッシュメモリを搭載したSSDを、2008年に世界で初めて<sup>(注1)</sup>開発した。その後、PC向けSSDはMLC NANDフラッシュメモリが主流になり、近年、スティックタイプSSDによる超薄型ノートPCや、低待機電力による長時間駆動ノートPCといった、HDDでは実現できなかった付加価値のあるノートPCの発展に貢献している。

Notebook PCs equipped with solid-state drives (SSDs), featuring shock and vibration durability due to the lack of moving parts, appeared on the market around 2006. However, single-level-cell (SLC) NAND flash memories were used in SSDs at that time, and it was believed that multilevel-cell (MLC) flash memories, while they could double the storage capacity compared with SLCs, were not applicable because of a lack of reliability.

In 2008, Toshiba developed the world's first SSD equipped with MLC NAND flash memories and realizing a low write application factor (WAF), thereby increasing the reliability of MLC SSDs. This was achieved through the analysis of actual PC users' workload in detail, as well as the development of a new NAND cache system and an appropriate NAND management algorithm. MLC NAND flash memories have now become the mainstream for SSDs used in notebook PCs, and are contributing to new values such as a thin and small form factor (SFF), very low standby current, and instant-on capability.

## 1 まえがき

NANDフラッシュメモリが1987年に開発されて以来、“HDD (ハードディスクドライブ) の置換え”は長年のテーマであった。近年、ノートPCにスティックタイプSSD (ソリッドステートドライブ) が搭載されるようになることで、この置換えが本格的に始まった。

1990年代、SSDはシリコンディスクやフラッシュドライブなどと呼ばれ、モータやピックアップなどの可動部品がないため耐振動性や耐衝撃性に優れているという長所があり、一部の産業機器などに採用されていた。しかし、NANDフラッシュメモリはHDDに比べて非常に高価であったため、採用されたのは限られた機器だけであった。

2000年代に入り、MLC (Multi Level Cell) NANDフラッシュメモリ (以下、MLC NANDと略記) が開発された。これにより、SDカードやシリコンオーディオをはじめとした、NAND

フラッシュメモリを使ったアプリケーションのビット当たりのコスト (ビットコスト) が大幅に低減され、MLC NANDの採用が急速に進んだ。しかしその一方で、MLC NANDは信頼性に懸念があったため、PCへは搭載できないと考えられていた。

このような状況のなか、東芝は、SSD開発のためのプロジェクトチームを2006年に設立した。このプロジェクトチームに課せられたチャレンジは、世界で初めてMLC NANDを搭載したSSD (以下、MLC SSDと略記) を開発することであった。このためには、最先端のMLC NANDの特性を理解してPC用途に耐えられる信頼性を確保すると同時に、HDDはもちろん、当時リリースされていたSLC (Single Level Cell) NANDフラッシュメモリ (以下、SLC NANDと略記) を搭載したSSD (以下、SLC SSDと略記) をも超えるパフォーマンスを達成する技術開発が必要であった。当社は、これらの課題を克服し、MLC SSDを2008年に世界で初めて開発した。

ここでは、MLC SSDを開発するにあたって克服した技術的課題と、そのために採用した技術について述べる。

(注1) 2008年2月時点、当社調べ。

## 2 MLC SSDの技術課題

MLC NANDの最大の特長は、ベースになるSLC NANDと比べて2倍の容量をサポートできることにある。これは、SSDにした場合、PC用途で必要だと考えられている64 Gバイトや128 Gバイトの製品のコストを、一気に1/2にする可能性を持つということである。

しかし、2倍の容量をサポートできる反面、MLC特有の、次のような課題があった。

- (1) 読出し、書込み速度の低下
- (2) 読出し時のエラービット増加
- (3) 消去ブロックサイズの増大
- (4) 書換え回数の低下

これらのなかでも、当時、MLC SSDがPC用途に使用できないと信じられていた最大の理由は、“書換え回数の低下”であった。一般的なPCの使用期間とされる3年間に、記憶素子であるMLC NANDの書換え回数が上限値を超えてしまい、保存してあるデータが消失してしまうことが懸念された。

MLC SSD開発において、最初に解決しなければならない課題は、この書換え回数の低下であった。当時のSLC NANDが通常10万回書換え可能だったのに対し、MLC NANDは3,000～5,000回程度と、1/20～1/30程度の書換え回数しか保証できず、この書換え回数の範囲で、MLC SSDが一般的なPCとしての使用条件に耐えることを証明する必要がある。

## 3 MLC SSDの寿命

MLC SSDの寿命は、搭載されるMLC NANDの寿命にはほぼ等しいが、実際の使用状態では、ホストシステムからSSDへ書き込まれるデータ量と、SSDに搭載されたNANDフラッシュメモリに書き込まれる実際のデータ量は一致しない。これは、NANDフラッシュメモリの特徴として、消去ブロックサイズが書込み単位より大きいことに起因しており、消去ブロックサイズ以下のデータの上書き保存時に、消去ブロックサイズ分の書込みが発生するためである。この書込み時の制約は“書込み効率”(WAF: Write Amplification Factor)と呼ばれる指標で、ホストシステムからの書込み量に対し、実際にNANDフラッシュメモリに書かれるデータ量が何倍になるかで表される。例えばユーザーが、消去ブロックサイズ2 MバイトのNANDフラッシュメモリに、編集中の200 Kバイトのファイルを上書き保存したとすると、特殊な場合を除いては、2 Mバイトの消去ブロックの消去と書換えが発生する。このため、ホストシステムからの200 Kバイトの書換え要求に対して、実際には2 Mバイトの書換えが発生することになり、書込み効率は10になる。したがって、3,000回書換え可能なMLC SSDであっても、このように書込み効率が10の使い方を毎回実行した場合は、ホ

ストシステムから見て、300回相当のデータ書込み量で、そのMLC SSDは寿命を迎えることになる。

したがって、MLC SSDがPCで使えることを証明するには、書込み効率を低く抑えるSSD制御アルゴリズムやアーキテクチャを設計する必要がある。そのためには、ユーザーが実際に使うデータのサイズ及び使う頻度と、1日当たりに保存するデータ量を正しく理解する必要がある。

## 4 ユーザーモデル

書込み効率を低く抑えるSSD制御アルゴリズムやアーキテクチャを設計するには、ユーザーの実際の使用状況をモデル化する必要がある。このために、家庭用PCよりも稼働時間が長く、データ書込み量も多いと思われる、当社内でのPCの使用状況をモニタしてモデル化を行った。延べ2か月にわたり、200名を超えるユーザーのアクセス状況をモニタした結果、書込みデータ量及びハイバネーション量<sup>(注2)</sup>ともに対数正規分布で近似できることがわかった。一例として、書込みデータ量の分布を図1に示す。データを正規分布にあてはめるとき、その平均 $\mu$ 及び標準偏差 $\sigma$ それぞれについて次の値が得られた。

- (1) 書込みデータ量

$$\mu \approx 2.2 \text{ Gバイト/日}$$

$$\sigma \approx 0.5 \text{ Gバイト/日}$$

- (2) ハイバネーション量(約1.05 Gバイト)

$$\mu \approx 1.2 \text{ 回/日}$$

$$\sigma \approx 1 \text{ 回/日}$$

したがって、当社内の平均的なユーザーは、PCに1日当たり約4 Gバイト(2.2 Gバイト+(1.05 Gバイト×1.2回)=3.46 Gバイト)のデータ量を書き込んでいることが判明した。

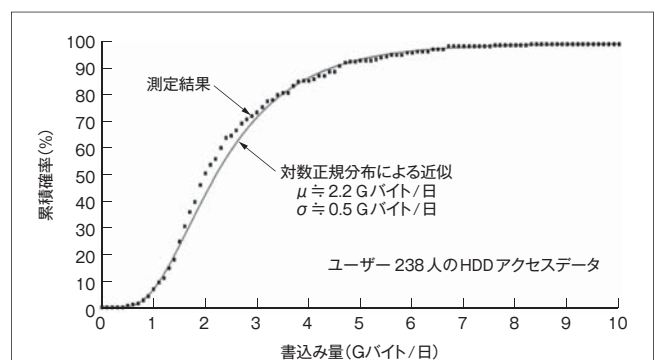


図1. PC使用状況の調査結果の例 — MLC SSDが実際にPCで使用可能な信頼性が確保できるかを実証するため、社内ユーザーのPCへの書込みデータ量を調査してモデル化した。

Example of result of monitoring amount of write data generated by PC users

(注2) ハイバネーションは、システム状態をSSDやHDDなどの外部記憶装置に保持してシステムを中断させること。ハイバネーション量は、その時に保持するデータ量。

## 5 ランダムアクセスとシーケンシャルアクセス

64 GバイトのMLC SSDに対して、1日約4 Gバイトのデータ書込み量は1/16の値であり、3年間毎日PCを使用したとしても書換え量は約70回に相当し、MLC NANDの書換え寿命3,000～5,000回に対して十分小さい値である。しかし、実際にMLC NANDが書き換えられる量は、この値に前述の書込み効率を掛けたものになる。書込み効率は、書き込まれるデータのサイズに強く依存し、512バイトや4 Kバイトなどの小さい単位でのランダムアクセスを伴う書込み時には、1,024のように非常に大きな値を取る。一方、同じ小さい単位での書込みでも、シーケンシャルアクセスであれば、書込み効率を低く抑えることが可能である。このため、モニタ結果である約4 Gバイトというデータ書込み量を構成する個々のデータのサイズ及びそれらのアクセス方法を理解することが、MLC SSDの寿命を決定する、適切なSSD制御アルゴリズムを設計するために重要になる。

そこで、実際のホストシステムからのアクセスを検討した。

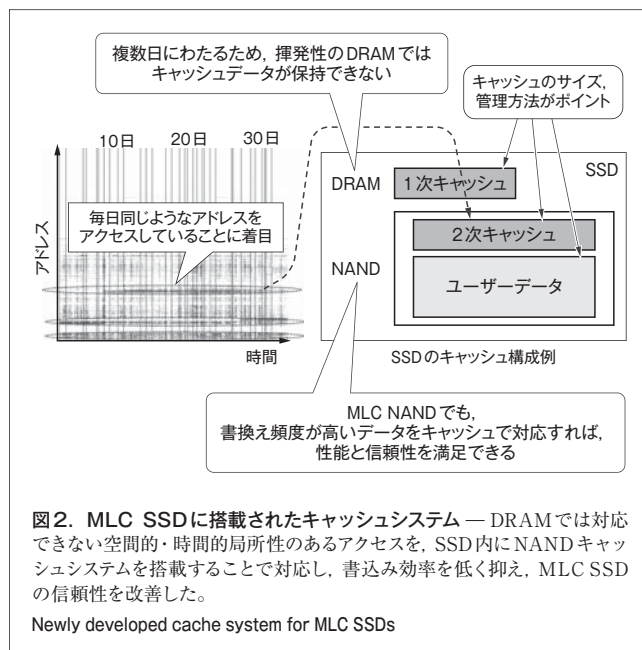
ユーザーモデルと同時に調査をした結果、全書込み量の約66%が、書込み効率が非常に大きな値になる4 Kバイトという最小書込みサイズであり、また、全体の90%の書込みが、消去ブロックサイズよりもはるかに小さい16 Kバイト以下であることがわかった。更に、ランダムアクセスも頻繁に行われていることがわかり、SSD制御アルゴリズムの設計を容易にするデータはまったく得られなかった。

そこで、MLC SSDの書込み効率を低く抑えるため、次章で述べる“キャッシュシステム”を新たに開発して適用することにした。

## 6 キャッシュシステム

小データのランダムアクセスが、ホストシステムからの基本的なアクセスパターンであることがわかったが、同時に、そのアクセスに関して時間的及び空間的な局所性があることもわかった。そこで当社は、キャッシュシステムを導入することで、小データのランダムアクセスを吸収し、書込み効率を低く抑える設計が可能であると考えた。しかし、その局所性は、数十MバイトのDRAMでカバーできるほど小さい領域ではなく、かつ、複数日にわたる時間的局所性があるため、揮発性のDRAMではキャッシュデータが保持できないという問題もあった。そこで、通常のDRAMキャッシュに加え、2次キャッシュとしてNANDそのものを利用し、専用のSSD内ファイルシステムを実装することによりこの問題を克服した(図2)。

SSD内NANDキャッシュシステムの特徴は、通常のDRAMキャッシュと異なり上書きができないため、更新が必要なデータは消去済み領域に書き込み、かつ、有効なデータだけを効



率良く集めてキャッシュに利用できる消去済み領域を確保するガーベージコレクションを実装していることである。このSSD特有のキャッシュシステムの開発によって、MLC SSDの寿命を決定する書込み効率は飛躍的に低減され、当社の場合は3～6程度の値を実現した。この結果、1日に4 Gバイトはもちろん、十数Gバイトの書込みがあった場合でも、一般的な使用方法であれば、MLC SSDが書換え寿命に達することがないように設計することができた。

## 7 SSDのパフォーマンス

消費者がSSDに期待する長所の多くは、HDDでは実現できない読出し・書込み速度の高速化であり、それらは、SSDに搭載されるNANDフラッシュメモリ単体の性能と、SSDの物理構成に依存する。SSDの物理構成は一般に、チャンネル数とバンク数の積として、4チャンネル×4バンクや、8チャンネル×4バンクなどのように表現される。

### 7.1 読出し性能

SSDの最高読出し性能は、NANDフラッシュメモリ単体のインタフェーススピード×チャンネル数で決まる。当社の最初の世代のSSDは、NANDフラッシュメモリ単体のインタフェーススピードが80 Mバイト/sで、構成が4チャンネルであったため、理論上の最高読出し性能は320 Mバイト/sであった。

近年、単体で133 Mバイト/sや200 Mバイト/sのインタフェーススピードを実現するトグルNANDフラッシュメモリが開発され、500 Mバイト/sを超えるSATA gen3 (第3世代 Serial Advanced Technology Attachment) SSDの実現に大きく貢献している。

## 7.2 書込み性能

最高性能はNANDフラッシュメモリ単体の書込み速度×並列書込みチップ数で決まる。標準的なMLC NAND単体の書込み速度を10 Mバイト/sとすると、4チャンネル×4バンク構成や8チャンネル×2バンク構成で、16チップ並列書込み動作をさせた場合、最高書込み性能は、10 Mバイト/s×16=160 Mバイト/sになる。消費電力を気にしなければ、8チャンネル×4バンクの構成とした32チップ並列動作で320 Mバイト/sの書き込み速度になり、SATA gen2の3 Gビット/sを超える性能を実現できる。

## 7.3 TRIMコマンド

前記の最高書込み速度を実現するには、必要なデータだけを消去済みのブロックに書き込む必要があるが、実際のSSD内部では、消去済みのブロックは少なく、また、有効なデータだけで埋め尽くされたブロックが少ないことも多い。このような状態のSSDに書込みを行うときには、複数のブロック内から有効なデータだけを集め、別のブロックに移動し、元のブロックを消去して書込みに備えるといった動作が必要になる。このため、実際の使用状態での書込み性能が低下し、極端な場合には、PCが一時的に応答しなくなったように見えることもある。

この問題を解決する一つの方法として提案されたのが、TRIMコマンドである。TRIMコマンドは、例えばファイルが消去された場合、消去されたファイルが記録されていたセクタが無効になったことをSSDに通知するコマンドである。TRIMコマンドを採用することで、無効になったデータを書込み時にコピーする必要がなくなり、速度低下が抑えられる効果が期待できる。ただし、このTRIMコマンドを受け取った後の動作はSSD側で決まるため、全てのSSDで効果があるとは限らない。

# 8 SSDの新しい価値

MLC NANDの採用によって高性能、高信頼性、及びコスト低減を達成したSSDだが、HDDに比べビットコストでは依然として10倍以上高価なデバイスである。このため今日でも、SSDが搭載されているのは一部のハイエンドノートPCに限られている。そのような状況下でも、当社は、以下に示すような、新たな価値の提案を続けている。

## 8.1 低待機電力と復帰時間

HDDにはスピンドルモータがあるために、アクティブアイドル時の待機電力は1 W弱となるものが多い。これに対してSSDの場合、待機電力を0.1 Wレベルまで引き下げることが可能になる。HDDの場合も、スリープ状態に入れることによって待機電力を0.1 Wのレベルまで削減できるが、復帰時間が延びてしまい、SSDのように待機状態から高速で復帰するこ

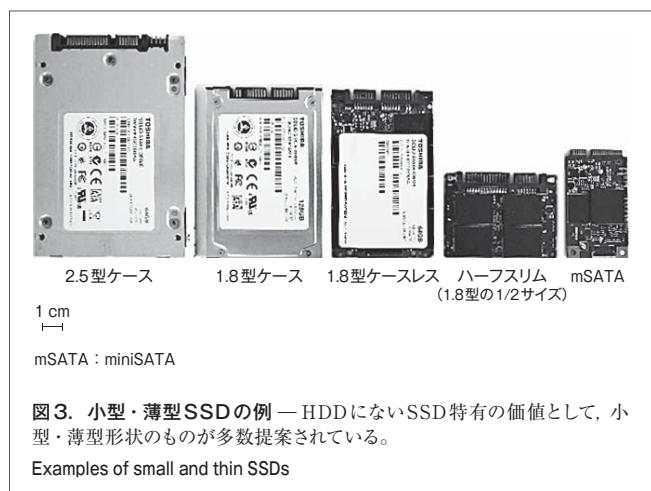
とはできない。このSSDの新しい価値を活用して、ノートPCのバッテリー駆動時間を飛躍的に伸ばし、HDD搭載のノートPCと差別化を行うことができる。このように、SSDを搭載することで、以前から求められていたインスタントオンPCを擬似的に実現できるようになった。

## 8.2 小型・薄型形状

開発当初のSSDは、HDDと互換性を持たせるために2.5型や1.8型のケースに収納していた。しかし近年、MLC NANDの大容量化や、NANDフラッシュメモリチップを8枚搭載したBGA (Ball Grid Array) パッケージの開発などに伴い、小型及び薄型でも、128 Gバイト、256 Gバイトといった大容量のSSDが実現できるようになった(図3)。この結果、ノートPCの内部で、ドライブに必要な体積が減り、その分、バッテリー容量を増やして長時間バッテリー駆動を可能にしたり、HDDでは不可能な、スティックタイプSSDを搭載した超薄型のスタイルリッシュなノートPCを実現できるようになった。

## 8.3 更なる高速ドライブの実現

高速なHDDとして7,200 rpmのHDDが知られているが、その読出し速度はSATA gen2が持つ3 Gビット/sで十分である。しかし、SSDでは、NANDフラッシュメモリのインタフェースの高速化と並列動作数を増やすことによって、SATA gen3が持つ6 Gビット/s以上の高速読出しも可能になる。このような特長から、昨今では、SAS (Serial Attached SCSI (Small Computer System Interface)) の12 Gビット/sや、PCI Express<sup>®</sup>(注3) (Peripheral Component Interconnect Express) 3.0の高速インタフェースをいち早く取り入れるデバイスとしてSSDが想定されている。今後、適切なインタフェースが開発されしだいSSDに搭載され、エンタープライズ向けを中心にサーバやストレージ機器の性能向上に貢献すると期待される。



(注3) PCI Expressは、PCI-SIGの登録商標。

## 9 SSDの更なる発展に向けて

クラウドコンピューティングの時代が到来し、SSD専用のノートPCの登場や、エンタープライズ向けサーバやストレージ機器への搭載という形で、今後SSDの市場が広がることはまちがいない。その発展をより確固なものにする技術として、RAID (Redundant Array of Independent (Inexpensive) Disks) に代表される信頼性向上のための機能が重要になる。これは、SSD固有の問題ではなく、HDDを使ったシステムでも過去に開発されてきたが、突発的な故障に対するデータ保護機能を、システムとして実装する方法が重要である。

SSDに搭載されるNANDフラッシュメモリは、年々微細化が進むにつれて物理的加工が難しくなり、ビットエラーをはじめ、書き込み単位でのエラーや消去ブロック単位でのエラーなど、一連の半導体特有のエラー頻度が高くなる。これらのエラー対策のために、BCH (Bose-Chaudhuri-Hocquenghem) 符号をはじめとするECC (Error Check and Correct) を強化して書き込み単位での信頼性を確保した。また、このRAID同様の対策をドライブ内で行うことによって、ページ単位、ブロック単位、及びチップ単位での突然のエラーに対処できるようになる。これによって、最先端のNANDフラッシュメモリを搭載し、より低いビットコストと、より高い信頼性を併せ持ったSSDを、ユーザーに提供することができる。

SSDの更なる発展は、クラウドコンピューティングの更なる発展へつながっている。

## 10 あとがき

ここでは、MLC NANDをベースにしたSSDの技術の概要について述べた。

SSDのインタフェースは、SATAやPCI Expressなどに限られたものではない。デジタルカメラなどで使用するSDカード、及びタブレットPCやスマートフォンに搭載されているe・MMC<sup>TM</sup>(注4)も、それぞれインタフェースは異なるが、NANDフラッシュメモリを使いこなすための課題や、基本的なNANDコントロール技術は、これまで説明してきたSSDと同じ考え方に基づいている。つまり、SSDの発展を支える技術は、NANDフラッシュメモリを使いこなす技術であり、それは、その他のNANDフラッシュメモリを使ったアプリケーションを支える技術とつながっている。

(注4) e・MMCは、JEDEC Solid State Technology Associationの商標。

更に将来的には、NANDではない新しいフラッシュメモリが登場するものと予想される。

この特集の他の論文でも触れられているとおり、新しいフラッシュメモリは、既に基礎研究が進められており、この新しいフラッシュメモリをドライブとして使いこなすための準備も始める必要がある。

しかし、現在提唱されている新しいフラッシュメモリの原理は複数あり、ドライブとして使いこなすための課題も多岐にわたる。それらを克服して、将来、クラウドコンピューティングの時代を支える、SDカード、e・MMC<sup>TM</sup>、及びSSDを開発して実用化するため、当社は、大学をはじめとする様々なパートナーとともに技術開発を進めていく。



土屋 憲司 TSUCHIYA Kenji

セミコンダクター&ストレージ社 ストレージプロダクツ事業部  
SSD応用技術部グループ長。SSDの事業戦略立案及び商品  
企画業務に従事。  
Storage Products Div.