

瓦記録技術を適用したデータセンター向け大容量HDD

Large-Capacity HDDs Applying SMR Technology for Data Centers

下村 和人 SHIMOMURA Kazuhito

近年、膨大な電子情報を保管するデータセンター向け大容量HDD（ハードディスクドライブ）は、情報インフラの発展には欠かせない製品として、需要がますます高まっている。一般のCMR（Conventional Magnetic Recording）方式のHDDを大容量化するには、磁気ヘッドや、磁気ディスク、信号処理回路といった主要部品の性能向上が不可欠となる。一方、屋根瓦をふくようにデータを記録トラックに重ね書きすることで大容量化を図る瓦記録（SMR：Shingled Magnetic Recording）方式の技術開発も進められている。

東芝デバイス&ストレージ（株）は、SMR方式の適用で大容量化を実現するHDDを開発している。今回、既存の3.5型CMR HDDにデータセンター向けSMR仕様のファームウェア（FW）を搭載し、磁気ヘッドと磁気ディスクの最適化を行うことで、記憶容量を約25％増大させた。

Large-capacity hard disk drives (HDDs) that allow data centers to store huge volumes of electronic information have become increasingly important as a key product for the development of information infrastructures in recent years. The two principal types of technical challenges to be addressed in order to increase the storage capacity of HDDs are (1) improvements in the performance of main parts including magnetic heads, magnetic disks, and the signal processing circuit in the case of conventional magnetic recording (CMR) HDDs; and (2) the overwriting of data on tracks, analogous to the shingling of a roof, in the case of shingled magnetic recording (SMR) HDDs.

Through the application of SMR technology and the optimization of magnetic heads and magnetic disks, Toshiba Electronic Devices and Storage Corporation is working toward realizing large-capacity HDDs for data centers. Experiments on a 3.5-inch CMR HDD product incorporating firmware appropriate for data writing in an SMR HDD have verified that it increases in storage capacity of approximately 25%.

1 まえがき

コンピューターの補助記憶装置として発展してきたHDDは、テレビ録画に代表されるパーソナルユースから膨大な電子情報を保管するデータセンター用途まで、様々な分野で使われている。また、データセンターに蓄積された電子情報は、ディープラーニングに代表されるAI技術の進歩に伴い2次活用も進んでいる。

これまでの20年、そして、今後の20年を見渡すと、データセンター向け大容量HDDは、情報インフラ社会の発展に欠くことのできない機器へと成長している。

こうした背景の中、SMR方式は、HDDの大容量化に適している反面、性能面では課題もあり、これを解決するためにホストシステム側の開発も進み始めている。

ここでは、東芝デバイス&ストレージ（株）が開発しているデータセンター向け大容量HDDに適用したSMR技術について、概要や今後の展望を含めて述べる⁽¹⁾。

2. CMR方式とSMR方式でのデータの記録と再生

HDDは、回転する磁気ディスク上で所定の半径位置（トラック）に磁気ヘッドを位置決めする。磁気ヘッドには、記録素子と再生素子が搭載されている。記録素子が、トラックの円周方向で所定の場所をS極とN極に磁化し、再生素子が、磁化の極性を読み出すことで、データの記録と再生を行う。

S極とN極の最小磁化反転間隔（ビットピッチ）とトラック間隔（トラックピッチ）を狭くし、磁気記録密度を上げることで磁気ディスク1枚当たりの記憶容量の増大を実現するのが、一般のCMR方式のHDD（以下、CMR HDDと略記）である（図1）。このCMR方式でトラックピッチを狭くするには、記録素子と再生素子のトラック方向（ディスク半径方向）の物理寸法を小さくする。この場合、小さくなった記録素子の記録能力や再生素子の再生能力は低下するため、磁気ヘッドと磁気ディスクの磁気特性の向上が必要となる。

一方、SMR方式のHDD（以下、SMR HDDと略記）は、記録進行方向に対して屋根瓦をふくようにデータを重ね書き

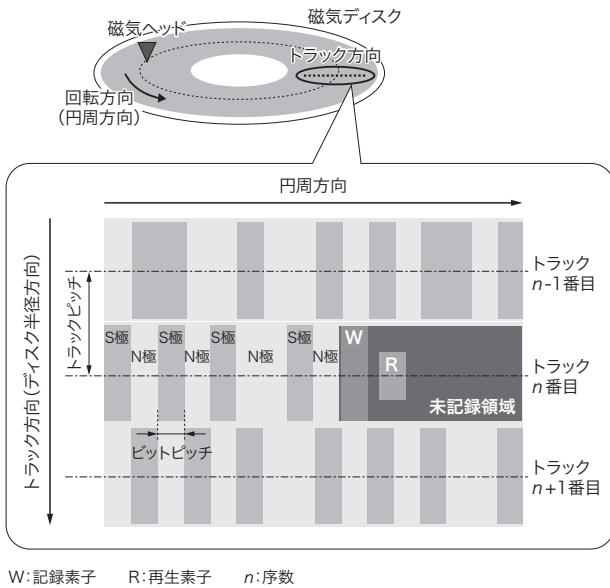


図1. HDDを構成する磁気ディスクと磁気ヘッドの概要

回転する磁気ディスク上のトラックに、記録素子と再生素子が搭載されている磁気ヘッドを位置決めして、データの記録や再生を行う。

Outline of head and media of HDD

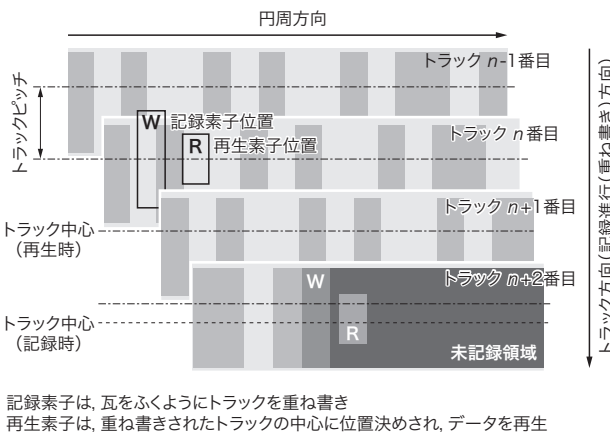


図2. SMR方式のトラックと、記録ヘッド、再生ヘッドの概略

屋根瓦をふくようなデータの重ね書きにより、トラックピッチを狭くできる。

Schematic diagram of data tracks, write head, and read head of SMR HDD

するため、記録素子などの部品構成が同じCMR HDDよりもトラックピッチを狭くできる利点がある（図2）。

磁気ディスク上のSMR領域は、バンドとガードバンドで構成されている。バンドは、トラックの重ね書きを連続して行うこと（シーケンシャルライト）でトラックピッチを狭くできる領域である。ガードバンドは、バンドの最後尾のトラックを記録したときに、隣接バンドの先頭トラックのデータを破壊しないようにするのが目的である（図3）。

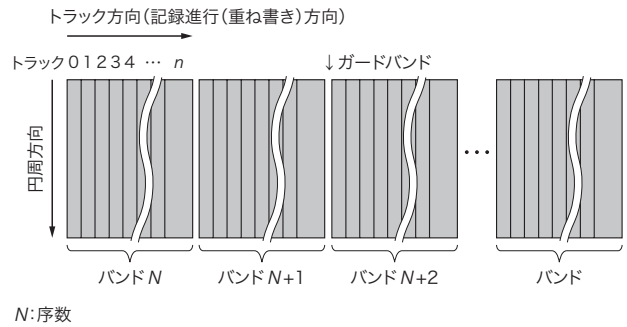


図3. SMR領域のトラックとバンドの構成

書き換え動作の効率化や仕様などにより、連続して重ね書きするトラックを規定する。ガードバンドは、隣接バンドの先頭トラックのデータを破壊しないようにするためである。

Configuration of shingled writing tracks and bands

3. データセンター向けSMR HDDのFW仕様

SMR方式のFWは、DM (Drive Managed) タイプとHM (Host Managed) タイプの二つに分類される。

DMタイプは、従来のHDDのコマンドで動作することを前提としており、CMR HDDの置き換えができる。ただし、ホストシステムのアクセス条件によっては、ランダムライト性能の変動が大きくなる。

HMタイプは、ホストマネージ用コマンドで動作することを前提としている。ホストシステムはSMR HDDの記録動作を最適化することで、安定したランダムライト性能を得ることができるが、ファイルシステムなどの開発をホストシステム側で行うが必要になる。

データセンターのHDDユーザーは、HMタイプの利用検討を進めており、HMタイプの専用コマンドは、情報技術規格国際委員会 (INCITS: International Committee for Information Technology) の下部組織である技術委員会で、SAS (Serial Attached SCSI (Small Computer System Interface)) 系を扱うT10と、SATA (Serial Advanced Technology Attachment) 系を扱うT13で規格化されている。

3.1 DMタイプのランダムライト動作と性能

DMタイプでは、ホストシステムから送られたデータは、HDDの内部動作によって、磁気ディスク上に設けられたメディアキャッシュ (MC) に一時記録される。MCの空き容量が減ってくると、データはMCからSMRバンドへ移動される。データの書き換えが必要なランダムライトでは、対象データが含まれるSMRバンドについて、全てのデータを読み出し、MCに一時記録された書き換え用のデータと入れ替えて、新SMRバンドとして旧SMRバンドとは物理的に異なる

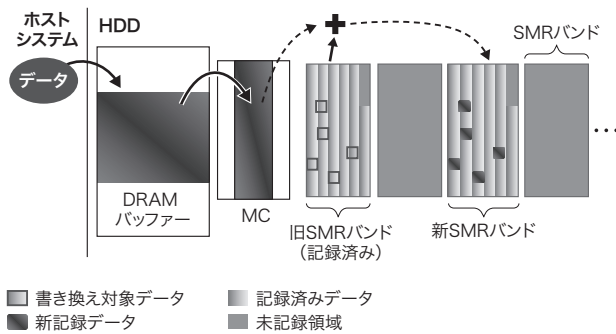
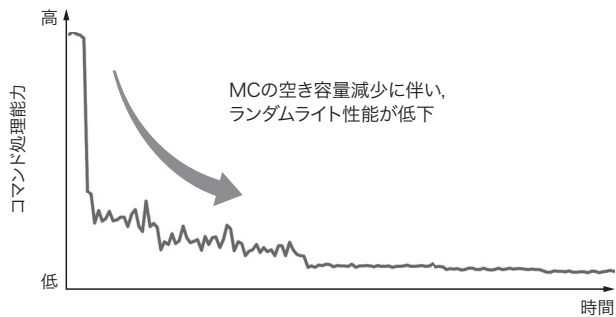


図4. DMタイプのランダムライト動作

MCを活用しながら、データの書き換えが必要なランダムライトなどの内部動作を行う。

Data rewriting procedure in drive-managed (DM) SMR HDD



MCの空き容量減少に伴い、ランダムライト性能が低下

*製品では、所定性能が一定時間維持できるように、対象アプリケーションごとに最適化し、急激な性能変化が起きないように調整されている

図5. DMタイプのランダムライト性能の例

MCからSMRバンドへデータを移動させる内部動作の頻度によって、性能が変動する。

Example of random data writing performance of DM SMR HDD

る場所にシーケンシャルライトする(図4)。

ランダムライトのデータは、必ずMCに一時記録されるため、短時間では高いランダムライト性能が得られる。しかし、ランダムライトの頻度が高くなり、MCの空き容量が少なくなると、MCからSMRバンドへデータを移動させる頻度が高くなるため、ランダムライト性能は低下する(図5)。

3.2 HMタイプのランダムライト動作と性能

HMタイプでは、MCを介さずに、WP (Write Pointer : SMRバンド内のデータの記録を始めることができるアドレスを示す指標) を基に、SMRバンドにデータを直接記録する。

WPは、各SMRバンドに一つ割り当てられている。データは、回転方向に対してWP以降の連続したアドレスにしか記録できないSMR方式の仕様になっている(図6)。

SMR領域には膨大な数のSMRバンドが存在するが、一度にアクセスできるSMRバンドの個数(最大SMRバンド数)はホストシステムが扱いやすいように仕様化されている。この

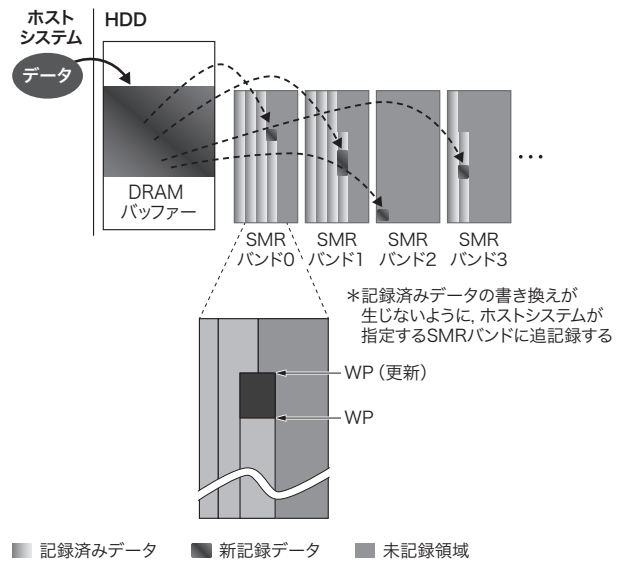
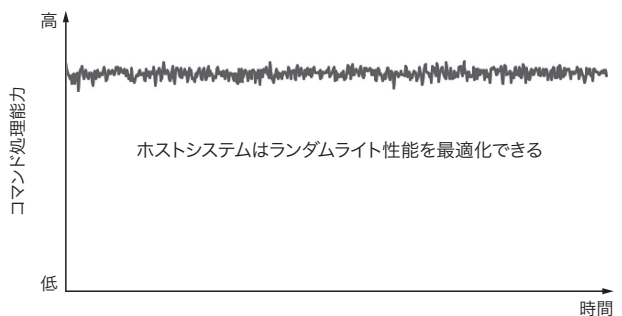


図6. HMタイプのランダムライト動作

ホストシステムは、MCを使用せず、WPを活用してSMRバンドにデータを直接記録する。

Data writing procedure in host-managed (HM) SMR HDD



ホストシステムはランダムライト性能を最適化できる

*一度に扱えるSMRバンドの個数(最大SMRバンド数)が仕様化されている(図は、128個のSMRバンドに対して、4,096バイトの固定サイズのライトコマンドをランダムに実行した例)

図7. HMタイプのランダムライト性能の例

ホストシステムは、HMタイプの仕様に基づき、所望のランダムライト性能が得られるように、HDDの記録動作を最適化できる。

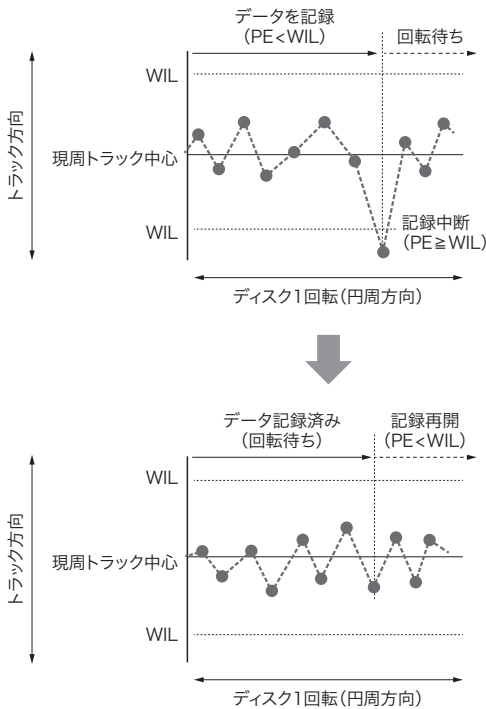
Example of random data writing performance of HM SMR HDD

ため、ホストシステムは、最大SMRバンド数の範囲でWPを活用し、所望のランダムライト性能が得られるようにHDDの記録動作を最適化できる(図7)。

したがって、ホストシステム側でHMタイプのSMR HDDを使いこなす技術を開発する必要がある。

4. SMR方式の特長を生かした品質確保技術

トラック中心に記録素子を位置決め制御しながらデータを記録するとき、HDDを搭載しているラックなどの外部振動要因でヘッドが揺れて、位置誤差(PE : Position Error)



●: 位置決め制御に関わるサンプリング時点のPE

*PEがWIL内に収まるまで、ディスクの回転待ちをするため、回転待ちが増えると、ライト性能は低下する

図8. ヘッドの位置誤差とデータ記録との関係

記録ヘッドの位置誤差を監視し、位置誤差が設計範囲を超えると記録を中断し、設計範囲に収まったら記録を再開することで、隣接トラックの記録済みデータへの悪影響を抑える。

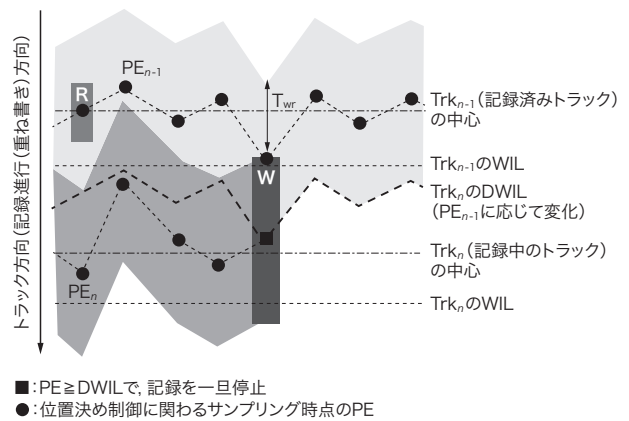
Relationship between head positioning errors and data writing operation

が大きくなることがある。PEが許容範囲（WIL：Write Inhibit Level）を超えると、隣接トラックの記録済みデータを上書きして壊すおそれがあるため、一旦記録動作を中断する。そして、ディスクの回転を待ち、PEがWIL内に収まったら、記録を再開する。CMR HDDでは、このようにして、隣接トラックの記録済みデータの品質を損なわないようにしている（図8）。

必要以上にWILを小さくすると、記録中断の回数が増加し、ライト性能が低下する。記録済みデータの品質確保とライト性能を両立させるWILの設計は、トラックピッチを狭くする上で重要である。

CMR HDDでは、固定値のWILを設定するが、シーケンシャルライトが特長のSMR HDDでは、記録済みトラック側と未記録トラック側で、異なるWILの設定をするDWIL（Dynamic Write Inhibit Level）を適用した。DWILの設定について、図9を用いて説明する。

データをシーケンシャルライトしているとき、そのトラッ



■: PE ≥ DWILで、記録を一旦停止

●: 位置決め制御に関わるサンプリング時点のPE

図9. DWILを適用したSMR HDDのデータ記録

データ記録の中断条件を動的に変えることで、隣接トラックの記録済みデータの品質確保と記録中のトラックの記録性能を両立させる。

Data writing of SMR HDD applying dynamic write inhibit level (DWIL)

ク Trk_{n-1} での記録ヘッドの位置誤差 PE_{n-1} をデータとして保存しておく。次のトラック Trk_n にデータを記録するとき、 Trk_{n-1} に記録されたデータの限界リード幅（ T_{wr} ：記録済みデータを再生できる最小幅）が確保できるように、 PE_{n-1} を基にDWILを設定する。DWILは、固定値のWILとは異なり、前周トラック Trk_{n-1} が現周トラック Trk_n に近づいていたときは小さい値を設定し、逆に遠ざかったときは大きい値になるように動的に設定する。

次に、記録中断について、以下で説明する。

現周トラック Trk_n を記録するとき、位置誤差 PE_n が未記録トラック Trk_{n+1} 側に大きくなり、WILを超える場合は記録動作を中断する。一方、位置誤差 PE_n が記録済みトラック Trk_{n-1} 側に大きくなったときは、DWILを超える場合だけ記録動作を中断する。

この動作により、トラックピッチを狭くしても、前周トラックの記録済みデータの品質を損なわずに、記録中断の頻度が高くなることを抑制できる⁽²⁾。

5. データセンター向けHDDへのSMR技術の適用

当社は、データセンターのHDDユーザーの要求に応えるため、既存の3.5型CMR HDDにDMタイプの2.5型HDD開発で培ったSMR技術を適用するとともに、HMタイプのFWを搭載し、磁気ヘッドと磁気ディスクの最適化を行った大容量HDDの開発を進めている（図10）。

磁気ディスク8枚で12 T（テラ： 10^{12} ）バイト及び9枚で14 TバイトのCMR HDDに対して、それぞれ、15 Tバイト及び17 Tバイトのデータセンター向けSMR HDDの製品開発を行っている。

+HMタイプ コマンド仕様	・ HMタイプのコマンド仕様をサポート
+SMR機能 (DMタイプ)	・ 必要なSMR機能の適用 ・ CMR HDD FWとのインターフェース
CMR HDD FW	・ 従来の3.5型CMR HDD向けFW

図 10. データセンター向けSMR HDDのFW構造

データセンターのHDDユーザーの要求に応えるため、SMR技術を適用するとともに、HMタイプのFWを搭載している。

Structure of firmware for SMR HDDs for data centers

6. 今後のSMR HDDの展望

CMR方式の良好なランダムアクセス性とSMR方式の大容量化の利点を1台のHDDで享受するため、アクセス頻度の高いデータをCMR方式で扱い、アクセス頻度の低いデータをSMR方式で扱う検討が始まっている。

INCITSでは、SMR領域とCMR領域を切り替えるコマンドの規格化を進めている。このコマンドを使うと、ホストシステムは、磁気ディスク上のSMR領域とCMR領域の割合を変えられることができるようになる。

当社は、次世代のSMR HDDとして、この技術への対応も進めている。

7. あとがき

当社のデータセンター向け大容量SMR HDD開発について、これまでのSMR技術の概要を交えて述べた。

SMR方式のコンセプトが生まれて10年が過ぎようとしている。磁気ヘッドや磁気ディスクの主要部品の性能向上だけに頼ることなく、HDDの記憶容量を増大できるSMR方式は、大容量化の視点で重要な技術となった。

HMタイプのコマンド仕様を通し、ホストシステム側の使いこなしの工夫も進み始めており、データセンターを中心にますますSMR HDDの需要は高くなると考えられる。

今後は、SMR方式に適した磁気ヘッドや、磁気ディスク、信号処理回路などの開発が進み、更なる大容量化につながるとともに、SMR HDDと親和性の高いシステムが開発され、性能面の課題が解決されていくと考えられる。

当社は、進化し続けるSMR方式をHDDに取り込み、データセンター向けHDDの大容量化を進めていくことで、サイバーフィジカルシステム(CPS)が支える社会変革の時代に貢献していく。

文 献

- (1) 下村和人. HDDの大容量化をけん引する瓦記録技術. 東芝レビュー, 2015, 70, 8, p.29-32.
- (2) 東芝, 河辺享之. ディスク記録装置, ディスク制御装置及び方法, 特許第5100861号, 2012-12-19.



下村 和人 SHIMOMURA Kazuhito
東芝デバイス&ストレージ(株)
ストレージプロダクツ事業部
HDD 製品技術部
Toshiba Electronic Devices & Storage Corp.