

ホワイト・ペーパー  
スマートストレージ・スタックベースHBAおよび  
RAIDソリューションでの  
SMRドライブの使用

2017年10月



## 目次

---

1 SMRドライブとは何か、そして利用される理由は？ .....	1
2 HBAまたはRAIDコンフィギュレーションにおけるSMRドライブ .....	3
3 実際の環境ではどのように利用されるか? .....	5
4 結論 .....	7

## 1 SMRドライブとは何か、そして利用される理由は？

瓦記録方式（shingled magnetic recording: SMR）とは、トラックの上に重ねて記録することにより、ハードディスクドライブ（HDD）の容量を増やす技術です。SMR HDDと区別するためにCMR（conventional magnetic recording）HDDとも呼ばれる通常のハードディスクドライブでは、ハードディスクドライブのトラック書き込みの最小幅は、データ書き込み上の信頼性を保つように設定される書き込み用ヘッドのサイズによって決まります。コンベンショナルHDDでは、トラックを重ねた記録は行いません。

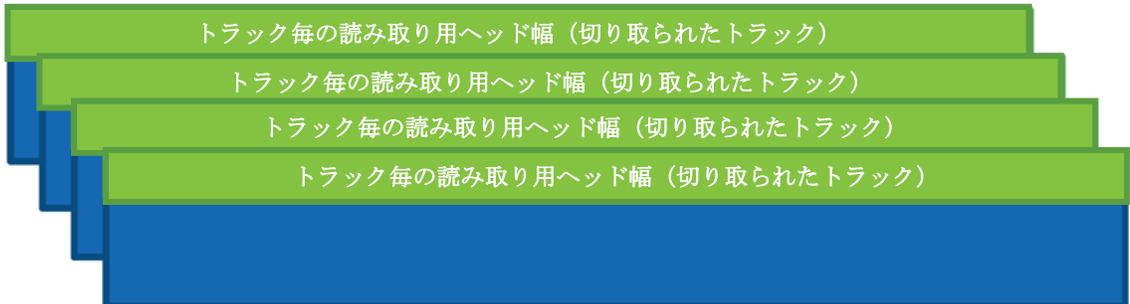
図1: コンベンショナルな書き込み（CMRドライブ、およびSMRのコンベンショナル・ゾーン）



SMRドライブは、データを読み取るために必要なトラックの幅と、HDDの読み取り用ヘッドの幅がかなり狭く、そのため書き込まれたトラックの全幅を必要とせずに確実にデータを読み取ることができる、という事実を利用しています。

以下の図は、SMRドライブがデータ密度を高めるために、書き込み中にどのようにトラックを重ね、使用可能容量を増やすかを示しています。これは、常にデータが終端に追加される限りにおいて機能します。データが中間に挿入されると、複数のトラックが同時に上書きされません。

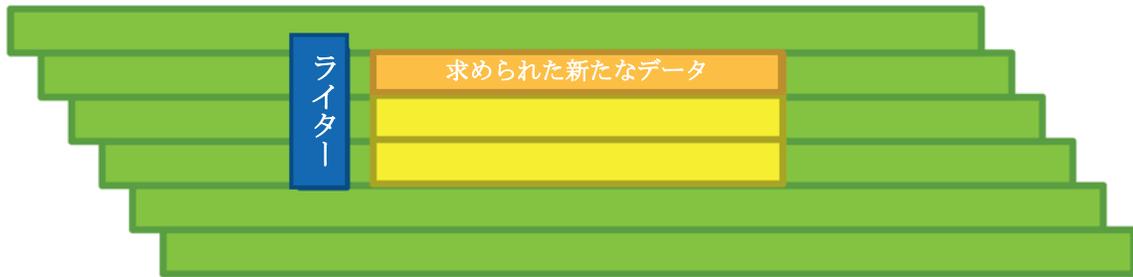
図2・重ねられたSMR書き込み



書き込み中のランダム性を許容するため、または複数のビデオストリームの記録や同時バックアップ実行のような、複数の連続ストリームからのデータを保存するため、各SMRドライブは、データを各バンドに追加することを可能にする書き込みポイントを有する複数のバンド（「ゾーン」と呼ばれる場合もある）を有しています。

SMRドライブの落とし穴は、次の図に示されるような、既存データの上書きです。これは、データ書き込みの際に、書き込み用ヘッドの幅が重なった複数のトラックに触れてしまうために起こります。データ損失を防ぐため、バンドのデータ全体をバッファに読み込ませてから、新たなデータをバッファ中の適切な場所に置く必要があります、それから最終的にバンド全体を再書き込みしなければなりません（読み取り、修正、書き込みのサイクル）。これは、潜在的に小さなブロックのランダムライト動作が、バンド全体の読み取り・書き込み動作に膨れ上がってしまう可能性を意味しています。

図3•SMRドライブ上での既存データの上書き



それに加え、バンドからデータを消去すると穴が残ります。このため、時々バンドを一掃するためのガベージコレクションを行い、バンドの終端に新たなデータを追加できるようにする必要があります。このプロセスは、発生する理由はやや異なりますが、フラッシュベースのSSDにおけるガベージコレクションとよく似ています。

データの上書きとガベージコレクションは、ヘッドの動きに数ミリ秒を要するその機構上の性質のため、ハードディスクドライブにおいて非常に時間を消費するプロセスです。このため、SMR HDDは、既にSSDのようなメモリーベースのストレージ機器の後塵を拝しているCMRドライブよりも、さらにランダムライト・ワークロードには適さないものになっています。

SMRドライブには、通常、トラックを重ねることのできないコンベンショナル・ゾーンがあり、これは、反応時間を改善するためのランダムライトデータのバッファ、またはさまざまなバンドの中の削除データもしくは書き込みポイント・ポジションなどを追跡するメタデータストレージとして利用されます。

複雑さが増すにもかかわらず、SMRドライブが採用されているのはなぜでしょうか？ その答えは、SMRドライブによって、コストを増やすことなくHDDの容量を増やすことができるためです。CMRドライブに使用されているものと同じ機構を活用し、SMR技術を利用することにより使用可能容量は15%~20%増加します。

フラッシュベースのSSDやその他の不揮発性メモリ（NVRAM）ベースのストレージなど、新たなストレージデバイス技術の導入によって、ハードディスクドライブがトランザクション・ワークロードに利用される頻度は減少しました。しかし、今後のデータ量の拡大が予測される中、I/OパフォーマンスよりもGB当たりの価格を最適化したコスト効率の良い大容量ストレージデバイスの需要は、いまだに増え続けています。

データは、ホットまたはコールドに分類されます。ホットデータとは、頻繁にアクセスされ、変更される頻度が高いもので、フラッシュベースのストレージやメモリに保存されます。ウォームデータとは、時々アクセスされるものの変更される可能性が低いものであり、コールドデータとは、アーカイブされ滅多に変更されないものです。コールドデータは、通常一度だけ書き込まれ、時の経過とともに、時々読み出されることがあります。コールドデータの一例としては、クラウドストレージにアーカイブされるフォトアルバムがあります。新しい写真は印刷される前に最終的な編集が行われますが、古いアルバムは変更されることなく、懐かしい思い出を振り返るために時々眺めるものです。

外付けUSB HDDや、アーカイブおよびバックアップ用のクラウドストレージは、SMRドライブの理想的なユースケースです。この記事を書いている時点で、既に数百万台のSMRドライブがUSBバックアップドライブとして採用されています。それは、より大きな容量を手頃な価格で提供しているためです。将来的に、SMRドライブは、おそらくアーカイブおよびコールドストレージに利用されるようになるでしょう。クラウドドライブや外付けUSBバックアップハードディスクドライブのさまざまな要求に対応するため、さまざまな種類のSMRドライブが採用されており、それらの種類と、その違いについては、次のセクションでご説明します。

## 2 HBAまたはRAIDコンフィギュレーションにおけるSMRドライブ

デバイスマネージド (DM) SMRドライブは、SMRを認識できる標準HBAに採用可能です。これらのドライブは、主にデスクトップまたはノートブック・システムの外付けUSBバックアップデバイスとして採用されています。

ホストアウェア (HA) SMRドライブに最適なHBAは、ZBC/ZACコマンドセットに対応したものです。これらのコマンドセットへの対応は、ホストマネージド (HM) SMRドライブには必須です。HAドライブは、ドライブの直接制御を必要とするアプリケーションによく使用されます。現在、OSおよびファイルシステム・ベンダーが、HMドライブ・ソリューションを開発中です。HMドライブ対応のための要件は、HMドライブがランダム性に全く対応していないという事実のため、他のドライブとは異なります。例えば、ほとんどの最新のオペレーティングシステムは、異なるサーバーCPUコア上で動作する複数のスレッドおよび複数のコマンドキューを使用して、システムのハードウェア・リソースを活用し、ハイエンド・メモリーベース・ストレージデバイスのI/Oオペレーションを最適化しています。そのようなアーキテクチャにおいては、全てのI/Oが順番に提示されるわけではなく、プロセスはHM SMRデバイスによってサポートされていません。

前述の例は、なぜSMRソリューションが、ハイパースケール・データセンターでの特定のアプリケーション専用化、特化する可能性が高いのかを示しています。これらのデータセンターは、HM SMR技術を最初に採用した利用者に含まれていますが、それは、それらが大量のストレージを配備し、ソフトウェアと全ての稼働アプリケーションを制御できるような安定した環境を必要とするためです。

リダンダント・アレイ・オブ・インディペンデント・ディスク (RAID) は、複数のストレージデバイスを、ホスト / オペレーティングシステムによって単一のストレージデバイスのように取り扱われるひとつの論理的ユニットに統合するものです。その目的は、複数のストレージデバイスを利用することによって、冗長性の追加と性能の拡張を実現し、データ可用性を向上することです。この目的のため、データはミラーリング (RAID1)、ストライピング (RAID0) することが可能であり、また可用性を高めるため、冗長情報を追加することができます (RAID5、RAID6)。

RAIDアダプタと呼ばれる専用のストレージアダプタは、不揮発性メモリによってRAIDアレイを管理し、また、これを使用することにより、ホストから冗長性計算をキャッシュおよびオフロードして、RAIDをさらに加速することが可能です。

RAIDコンフィギュレーションは、固定パターンを使用してRAIDアレイ中のストレージデバイスの中でデータを分散する方法を決定しており、そのため全ての書き込みデータが連続している必要があるHM SMRドライブには適していません (ライト・イン・プレイスに対応していない)。RAIDは、ドライブマネージドモード対応のDM SMRドライブおよびHAドライブとともに使用することが可能です。

RAIDは、SMRドライブとCMRドライブのいずれとも互換性がありますが、性能特性が非常に異なるため、これらの2つのドライブを同じRAIDアレイの中で混用することはお勧めできません。「鎖は一番弱い部分よりも強くはなれない」ということわざがあります。これと同じように、SMRドライブとCMRドライブを混用したRAIDアレイの性能も、SMRのみのRAIDアレイに近いものとなるのです。SMRドライブは、追加されたその複雑性のため、処理できるIOPSの数値が限定され、ランダムライト・ワークロード中のI/O要求に応答する際のレイテンシが一貫しないという問題があります。SMRドライブをRAIDアレイに組み込んでも、この実態は変わりません。

要するに、RAIDアレイ中のSMRドライブにも、個々のSMRドライブと同様の制限があるということです。しかし、RAIDコンフィギュレーションは、CMRドライブに対する場合と同様に、複数のSMRドライブの性能の集約を促進することが可能です。その結果、RAIDがより高度なデータ可用性を提供する一方で、ワークロードにおいて全体的に高いレベルの性能を達成することが可能です。

RAIDアレイ中のSMRドライブに対応するために、RAIDソフトウェアと不揮発性キャッシュがデータのバッファに役立ちます。しかし、それによってSMRドライブの特性を変えることはできません。それは、RAIDアダプタが、コマンドをバッファし、より緩和されたタイムアウトを提

供するという点において、できることには限界があるということです。このようなRAIDアレイに使用されるSMRドライブは、ごく限られたランダムライト性能が要求されるアプリケーションおよび環境においてのみ使用されるべきです。このルールは、HBAに接続される個々のSMRドライブにも当てはまります。そうしないと、ディスクとアダプタのバッファはすぐに容量を超えてしまい、I/O要求が適時に処理されないため、アプリケーションがスターベーションを起こしてしまいます。

間違った使い方をすれば、DM SMR HDDベースのRAIDアレイのI/O処理時間は、OSよりも大幅に長期化し、個々のSMRドライブのように、ハングやデータ損失さえも引き起こします。このため、RAIDでのDM/HA SMRドライブの使用は、ランダムライトI/Oのレベルが低く、かつ制御可能な環境においてのみ推奨されます。

### 3 実際の環境ではどのように利用されるか?

RAIDアダプタとHBAのいずれを適切なSMRドライブと動作させる場合にも、最適化が必要になります。前述した推奨動作モードへのSMR対応は、Microsemi Adaptec® HBA 1100 Series、martHBA 2100 SeriesおよびSmartRAID 3100 Seriesで利用可能です。

ユーザーがCMRドライブと見た目と動作が類似したDMおよびHA SMRドライブを識別することを支援するため、Microsemi Smart Storageアダプタは、ドライバ・タイプをレポートするデバイスの識別子を確認します。それがSMRドライブとして識別された場合、アダプタは、ユーザーにそれがスタンダードCMR HDDではないことを通知する、個別のアイコンによって表示します。以下のイラストは、それらのアイコンとその意味を示しています。

#### 図4・Enterpriseビューのアイコン

##### Enterpriseビューのアイコンの意味

アイコン	説明
	コントローラと直接接続されたディスクドライブまたはエンクロージャを備えたシステム
	コントローラ
	エンクロージャ
	論理デバイス
	maxCache デバイス（異常なし）
	ハードディスクドライブ
	ソリッドステートドライブ（SSD）
	SMRドライブ
	コネクタまたはその他の物理ドライブ

これらの通知は、デバイスタイプを明らかにし、ユーザーが同じRAIDアレイまたは論理ボリューム中でCMRドライブとSMRドライブを混用することを防ぐように設計されています。ユーザーがこのようなデバイスを混用したRAIDアレイを作成しようとした場合、アダプタは警告メッセージを表示し、RAIDまたは論理ドライブ（LD）の設定を防ぎます。HBAも、アイコンを使用してSMRドライブがHAかHMかを表示します。

このような情報を備えることにより、ユーザーは全体の容量に対してSMRのみの論理ボリュームとRAIDアレイを作成し、性能、容量、およびデータ可用性を向上することが可能になります（確立されたSMR技術の制限の範囲内で）。CMRドライブに対応したものと同一RAIDレベルが、論理ボリュームについて類似したデバイス数を持つSMRデバイスおよびRAIDアレイの双方においても利用可能です。

OSブートでさえも、最新のDM SMRドライブによって対応されています。読み取りと限られたランダムライトのみ行えばよいからです。しかし、まだSMRドライブが性能に影響するシナリオが存在しますので、このユースケースはお勧めできません。SMRドライブが激しいランダムライトアクセスを受けると、ドライブのランダムライトゾーンに汚染データが蓄積し、最終的にデバイスが反応しなくなります。このシナリオを回避するには、ドライブをSMRに適した環境で使用するようユーザーが徹底するしかありません。HAモードでのHM SMRドライブおよびHA SMRドライブの使用は、ローデバイスに限定されます。これらは、HBA 1100、SmartHBA 2100、および

SmartRAID 3100製品を含むMicrosemi Smart Storageアダプタによってサポートされています。Microsemiは、そのアダプタによるさまざまなHA SMR HDDの検証に成功しており、現在はランダムライトを受容できないために既存デバイスでは広く対応されていないHMデバイスへの完全な対応を実施するため、オペレーティングシステムおよびソフトウェア・ベンダーと協力しています。

SMRドライブの互換性およびサポートに関する詳細情報につきましては、Microsemiの販売代理店までご連絡いただくか、Microsemiのウェブサイトをご覧ください。

## 4 結論

---

SMRは、コンベンショナルHDDでは不可能な、重ねられたトラックへの書き込みを行うことによってHDDの容量を増やします。RAIDコンフィギュレーションは、固定パターンを使用してRAIDアレイ中のストレージデバイスの中でデータを分散するため、ライトデータが連続している必要があるHM SMRドライブには適しません。しかし、RAIDは、ドライブマネージドモードに対応したDM SMRドライブおよびHA SMRドライブには、使用することができます。RAIDアレイ中のSMRドライブには、個々のドライブとして使用されるSMRドライブと同様の制限がありますこの事実にもかかわらず、RAIDはCMRドライブに対する場合と同様に複数のSMRドライブの性能の集約を促進し、ワークロードにおける全体の性能レベルとデータ可用性の向上を実現することができます。

Microsemi Adaptec HBA 1100、SmartHBA 2100およびSmartRAID 3100シリーズの発売を始めとして、Microsemiのストレージ・アダプタ・ソリューションは、HM SMRまたはHA SMRドライブに必要なコマンド拡張をサポートします：ゾーン・ブロック・コマンドまたはゾーン・デバイスATAコマンドセット（ZBC/ZAC）。HAモードでのHM SMRおよびHA SMRドライブの稼働はローデバイスのみに限り、正常な稼働を確保するためにはオペレーティングシステム、ファイルシステムまたはアプリケーションのサポートが必要となります。

Microsemi製品に関する詳細については、[www.microsemi.com](http://www.microsemi.com)をご覧ください。



**Microsemi Corporate 本社**

One Enterprise, Aliso Viejo,  
CA 92656 USA  
アメリカ国内 : +1 (800) 713-4113  
アメリカ国外 : +1 (949) 380-6100  
Fax: +1 (949) 215-4996  
Eメール: sales.support@microsemi.com  
www.microsemi.com

© Microsemi Corporation. 無断複写・転載を禁ず。MicrosemiおよびMicrosemiのロゴはMicrosemi Corporationの商標です。その他の商標およびサービスマークの権利はすべて、それぞれの所有者に帰属します。

Microsemiは、本書の内容および特定用途に対する製品およびサービスの適合性に関して一切の保証も表明も行わず、かつ、製品もしくは回路の応用または使用に起因する法的責任も一切負担しません。この記載に従って販売された製品およびMicrosemiが販売したその他の製品は、限定された試験を受けており、ミッションクリティカル機器またはアプリケーションと併用すべきではありません。性能仕様の信頼性は高いと見なされていますが立証されてはいないため、購入者は購入した製品単品およびそれを搭載した最終製品の状態で、性能試験およびその他の試験をすべて実施し完了する必要があります。購入者は、Microsemiが提供するデータ、性能仕様、パラメータに依存しないでください。製品の適合性を独自に判断し、その製品の試験および検証を行うことは、購入者の責任です。Microsemiがこの記載に従って提供する情報は、「現状のまま」、一切購入者の責任で提供されるものであり、当情報に関する全体的なリスクは完全に購入者が負担します。Microsemiはいかなる関係者に対しても、特許権、ライセンス、およびその他のIP権利を、上記情報自体または上記情報が説明する内容との関連性の有無にかかわらず、かつ明示的にも黙示的にも付与しません。本書にて提供される情報の所有権はMicrosemiにあり、Microsemiは、予告することなくいつでも、本書の情報およびあらゆる製品やサービスに変更を加えることができる権利を留保しています。

Microsemi Corporation (Nasdaq: MSCC) は、航空宇宙、防衛、通信、データセンタ、産業市場向けの多岐に渡る半導体・システムソリューションを提供しています。取扱製品は、高性能の放射線強化アナログ混合信号集積回路、FPGA、SoC、ASIC、電源管理製品、世界の時刻基準を設定できる時間設定・同期デバイスと高精度時間ソリューション、音声処理装置、RFソリューション、デバイスクリート部品、企業向けの記憶・通信ソリューション、セキュリティ技術と拡張可能な改ざん防止製品、イーサネットソリューション、Power-over-Ethernet ICとミッドスパンのほか、カスタムデザインの機能とサービスなどです。本社をカリフォルニア州アリソン・ビエホに構え、世界中の拠点を含めた総従業員数は約4,800名になります。詳細は [www.microsemi.com](http://www.microsemi.com) をご覧ください

ESC-2171613