

テクノロジー ペーパー

# RAID復旧による ダウンタイムの削減

## はじめに

選択された特定のRAIDレベルに応じて、複数のディスク・ドライブのアレイを構成することで、データの信頼性、I/Oパフォーマンス、あるいはこれらの両方を最大限に高めることができます。

各RAIDレベルでは、こうした目的を達成するために、独自のテクノロジーの組み合わせ（ミラーリング、ビット・レベル、バイト・レベル、またはブロック・レベルのストライピング、専用または分散パリティ）が採用されていますが、RAIDソリューションには、「ドライブが故障した場合のRAID復旧には時間がかかり、エラーが発生しやすい」という共通の問題点が存在します。

近年、ストレージ・システムのスループット向上をはるかにしのぐペースでディスク・ドライブの容量が増大していることから、RAIDソリューションへの需要が高まっており、その結果として、RAID復旧にさらに時間がかかるようになっていきます。企業のRAIDセットを復旧させるには、数時間あるいは数日かかる場合もあり、長時間の復旧プロセス中にセカンダリ・ドライブが故障した場合などはさらに厄介な問題が発生します。このような状況下では、RAIDのデータは読取り不能となり、IT管理者はこのデータのバックアップ・バージョンを再処理しなければならなくなります。

従来のRAID復旧では、使用可能な他のドライブのパリティ・データを使用して、故障したドライブのデータを復元し、ホット・スペアに書き込むという複雑で時間のかかるプロセスを用いるのが一般的でした。シーゲイトは、シーゲイト RAID Rebuild™ テクノロジーの部分的障害コピー機能を用いて、より優れたソリューションを提供します。この機能を使用すれば、ホストはRAID復旧によるデータ回復を行う前に、故障したドライブからできるだけ多くのデータを迅速に取り出すことができるため、再作成が必要となるデータの量をずっと抑えることができます。この種のRAID復旧は、従来の方法よりも短時間で完了できる上、エラーの発生率も低くなります。

## シーゲイトRAID Rebuild™ テクノロジーの利点

シーゲイトでは、幅広い調査に基づき、RAID復旧の最も重要な基準として以下を特定しました。

- RAID復旧作業中の二次的な障害の発生リスク/影響
- RAID復旧作業中に発生するシステム・レベルでのパフォーマンス低下
- RAID復旧作業の開始から終了までにかかる時間

# RAID復旧による ダウンタイムの削減



RAID復旧プロセスでは、復旧に伴い時間のかかる読み込み/書き込み作業が必要となるため、ディスク・ドライブに多大な負担がかかります。現代の企業が有するドライブには膨大な量のデータが保存されているため、復旧に要する時間はどうしても長くなってしまいます。そして言うまでもなく、RAID復旧作業中にセカンダリ・ディスクに障害が発生するという事態も共通の懸念となっています。

さらに、こうした時間のかかる復旧作業はシステム・レベルでのパフォーマンス低下を招くため、RAIDで動作中のドライブであっても、保存データにスムーズにアクセスできないという問題も出てきます。RAIDアレイのホストI/Oよりも復旧プロセスを優先させることで、復旧に要する時間を短縮する（また、セカンダリ・ドライブの障害発生率を低くする）ことはできますが、この場合、システム・レベルでのパフォーマンスがさらに低下することになります。

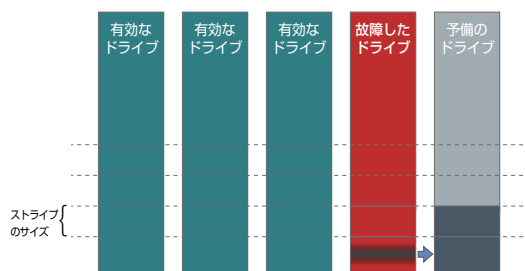
シーゲイトRAID Rebuildテクノロジーが搭載されているディスク・ドライブの部分的障害コピー機能は、上記のような問題を直ちに解決します。シーゲイトRAID Rebuildテクノロジーは、RAID復旧を実行する前に、故障したドライブからホストを通じてできるだけ多くのデータを取り出すことで、以下を実現します。

- 迅速でエラーの少ないRAID復旧
- システム・パフォーマンスへの影響の軽減
- 回復したデータへの迅速なアクセス

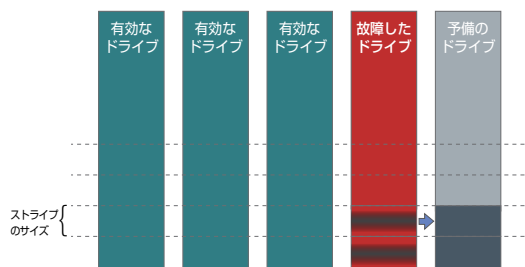
## シーゲイトRAID Rebuild™ テクノロジーの仕組み

### PIドライブのフォーマット

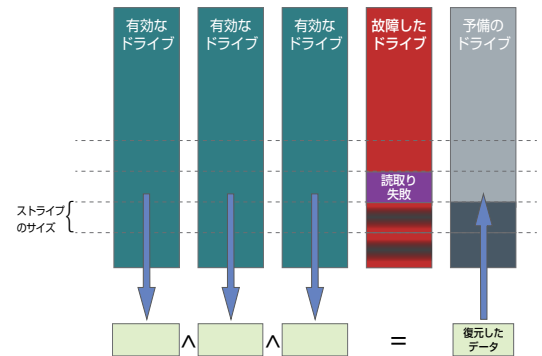
1. まず、故障したドライブからの読み取りを試みます。
2. 故障したドライブからコピーすることで、最初のストライプの復元を行います。



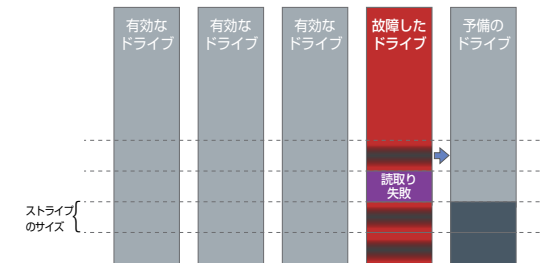
3. 故障したドライブからコピーすることで、2番目のストライプの復元を行います。



4. 故障したドライブからの3番目のストライプでエラーが発生した場合、有効なドライブを用いてデータの復元が行われます。



5. 故障したドライブからコピーすることで、4番目のストライプの復元を行います。



# RAID復旧による ダウンタイムの削減



## シーゲイトRAID Rebuild™テクノロジーが搭載されたSASドライブ

ホストが診断送信コマンドを発行することで、故障したドライブの特徴を示すよう要求します。ドライブが回転しない場合、有効なヘッド上で回転を試みます。それでもドライブが回転しない場合は、「HARDWARE ERROR」（ハードウェアの不具合）が表示され、シーゲイトRAID Rebuildの部分的コピー機能によるデータの取り出しを行うことができなくなります。

ドライブが上手く回転した場合は、部分的コピー機能の準備として、不要なバックグラウンド・アクティビティを排除し、使用不能なヘッドが含まれていないかどうかを判断して、メディアの書き込み保護を行い、エラー回復を無償の再試行に限定する特殊モードへと移行します。このモードは、ドライブの電源を一旦切って入れ直すまで有効となり、バス・リセットのプロセスを通じて継続します。

その後、入手可能なデータを取り出すため、ホストを通じてシーケンシャル読取り作業負荷が発行されます。このようなシーケンシャル読取りプロセスでは、再試行の割合に基づいて、使用不能なヘッドのリストが追加される場合があります。使用不能なヘッドにリストアップされているヘッドがコマンドに含まれている場合、問題のあるLBAはどこに存在するのか、また次に使用可能なLBAはどこに存在するのかを示すセンス・データがドライブに表示されます。この場合、ホストは次に使用可能なLBAで、シーケンシャル読取りプロセスを再開します。（ホストが待機読取りコマンドを発行している場合は、コマンドの一部が失敗し、次に使用可能なLBAへと移行することがあります。）

ホストは、依然として再構築が必要なLBA、すなわちドライブから読み取ることができなかったLBAのリストを保持します。<sup>1</sup>

## シーゲイトRAID Rebuild™テクノロジーが搭載されたSATAドライブ

ホストが、故障したドライブに対して、S.M.A.R.T. オフライン即時コマンドを発行します（ドライブが回転しない場合は、シーゲイトRAID Rebuildの部分的コピー機能でデータの取り出しを行うことができません）。このコマンドが発行されると、ドライブは不要なバックグラウンド・アクティビティを排除し、使用不能なヘッドが含まれていないかどうかを判断して、メディアの書き込み保護を行い、エラー回復を無償の再試行に限定する特殊モードへと移行します。ドライブの電源を一旦切って入れ直すまで、このモードは有効となります。

その後、読取りFPDMA待機コマンドを使って入手可能なデータを取り出すため、ホストを通じてシーケンシャル読取り作業負荷が発行されます。このようなシーケンシャル読取りプロセスでは、再試行の割合に基づいて、使用不能なヘッドのリストが追加される場合があります。使用不能なヘッドにリストアップされているヘッドがコマンドに含まれている場合、対応するステータス値およびエラー値がドライブに表示され、その後、次の手順へと進むために、ホストがログ・ページ0x10から問題のあるLBAと次に使用可能なLBAの読込みを行います（SATAプロトコルでNCQエラーが発生すると、ホストがこのログの読取りを行うまでは、デバイスで新しいコマンドを受理できなくなります）。ホストはその後、次に使用可能なLBAでシーケンシャル読取り作業を再開します。

ホストは、依然として再構築が必要なLBA、すなわちドライブから読み取ることができなかったLBAのリストを保持します。<sup>2</sup>

<sup>1</sup> SASドライブでは、コマンドごとに完全なエラー回復を再有効化し、使用不能なヘッドのリストを無効化するために、読取りコマンドで強制ユニット・アクセス (FUA) ビットが使用される場合があります。

<sup>2</sup> SATAドライブでは、コマンドごとに完全なエラー回復を再有効化し、使用不能なヘッドのリストを無効化するために、読取りFPDMA待機コマンドで強制ユニット・アクセス (FUA) ビットが使用される場合があります。

# RAID復旧による ダウンタイムの削減



## 結論

シーゲイトRAID Rebuild™ 部分的障害コピー機能が搭載されたディスク・ドライブには、非常に明確で説得力のあるメリットがあります。それは、故障したドライブから直接取り出せるデータの量が多いほど、RAID復旧作業で回復しなければならぬデータの量を減らせるということです。RAID復旧作業には、エラーが発生しやすく、時間がかかるという問題があります。

シーゲイトRAID Rebuildテクノロジーが実現する迅速な復旧により、セカンダリ・ドライブの故障の発生リスク/影響を最小限に抑え、RAID環境でデータ保全性を保護することが可能となります。このシーゲイト独自の特種なテクノロジーにより、システム・レベルでのパフォーマンス低下を軽減し、故障したドライブのデータを使用可能な状態へとすばやく戻すことができます。

## 加盟団体

主要な業界標準化団体の加盟企業として、シーゲイトは、T10<sup>3</sup> (SAS案：11-298) およびSATA-IO<sup>4</sup> (SATA案：SATA31\_TPR\_D144) に対し、両委員会が公表する標準仕様にRebuild Assistの名前でRAID Rebuild™ 機能を掲載してもらうため、オープン・スタンダード案を提出しました。

<sup>3</sup> T10は情報技術標準国際委員会 (INCITS) であり、米国規格協会 (ANSI) に認可され、同協会が承認した規則の下で運営されています。これらの規則は、各業界グループの合意によって自発的な規格が開発されることを意図したものです。INCITSが情報処理システム基準を策定し、ANSIがその策定・公表プロセスの認可を行っています。ANSIはまた、国際標準化機構 (ISO) および国際電気標準会議 (IEC) の合同技術委員会 -1 (JTC-1) の米国代表としての役割も果たしています。詳細については、<http://www.t10.org/> をご覧ください。

<sup>4</sup> シリアルATA国際機関 (SATA-IO) は、業界の大手企業によって開発された独立した非営利団体です。SATA-IOは業界に対し、SATA仕様の導入に関するガイダンスおよびサポートを提供しています。標準化されたSATA仕様により、これまで15年間に渡って使用されてきた従来の技術に代わり、今後最長で10年間のサポートを可能とする高速シリアル・バスが登場しました。SATA-IOの加入メンバーは、SATA仕様の開発に対して影響を与え、直接的に貢献する権限を有しています。詳細については、<http://www.sata-io.org/> をご覧ください。

[www.seagate.co.jp](http://www.seagate.co.jp)



南北アメリカ Seagate Technology LLC 10200 South De Anza Boulevard, Cupertino, California 95014, United States, +1 408 658 1000  
アジア/太平洋 Seagate Singapore International Headquarters Pte. Ltd. 7000 Ang Mo Kio Avenue 5, Singapore 569877, +65 6485 3888  
ヨーロッパ、中近東、およびアフリカ Seagate Technology SAS 16-18 rue du Dôme, 92100 Boulogne-Billancourt, France, +33 1 41 86 10 00

© 2011 Seagate Technology LLC. All rights reserved. Printed in USA. Seagate、Seagate Technology、およびWaveのロゴは、米国およびその他の国々におけるSeagate Technology LLCの登録商標です。Seagate RAID RebuildおよびRAID Rebuildは、米国およびその他の国々におけるSeagate Technology LLCまたはその関連会社の商標または登録商標です。その他の商標または登録商標は各社の所有物です。報告される情報は、最終版T10/SATA-IO基準に基づき、変更される場合があります。製品内容または仕様は、予告なく変更される場合があります。予めご了承ください。TP620.1-1110JP、2011年10月