

ホワイトペーパー

Fujitsu PRIMERGY サーバ

512e HDD : テクノロジー、パフォーマンス、構成

物理 4K セクターを使用した HDD の新技術により、ストレージ密度、およびエラー訂正機能が向上されています。その新技術の HDD が現在、PRIMERGY サーバに搭載されています。これは、外部に対して 512 バイトのセクターサイズにエミュレートした形になっており、弊社では 512e HDD と呼びます。最適なパフォーマンスを得るには、適切な環境にて使用する必要があります。

バージョン

1.2a_J_6



目次

ドキュメントの履歴	2
ハードディスクの動作説明	2
パフォーマンス	3
注意が必要なシステム構成	5

ドキュメントの履歴

バージョン : 1.2a_J_5 (2015-1-22)

初版

バージョン : 1.2a_J_6 (2015-6-23)

「注意が必要なシステム構成」を更新

ハードディスクの動作説明

長期にわたり、HDD の物理セクターの構成は、各セクターに 512 バイトのユーザーデータと一定のオーバーヘッド（「エラー訂正コード」など）を含むものでした。これらの HDD は、略して「512n」で呼ばれることもあります（「n」は「ネイティブ」を意味する）

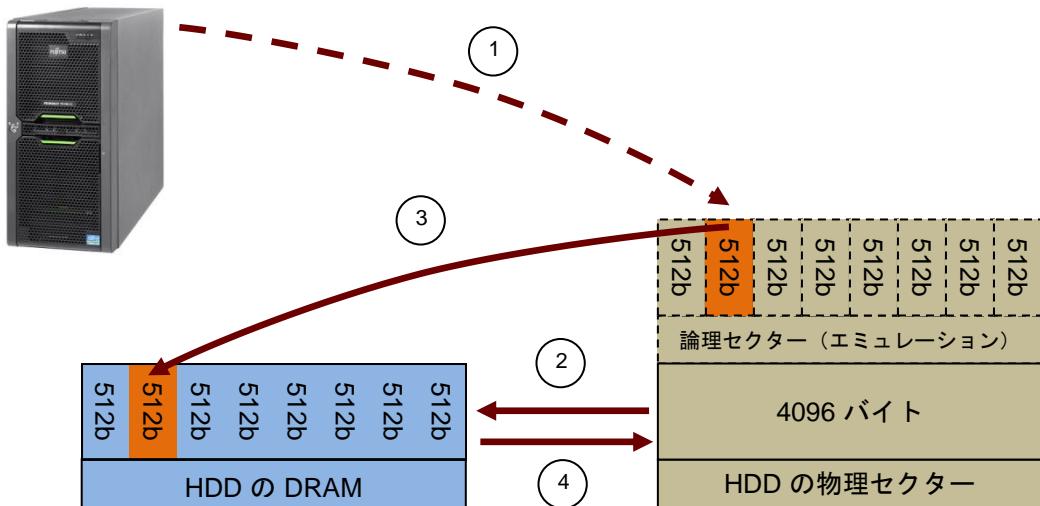
しかし、このセクターフォーマットは近年、その限界に達しています。今日要求されるストレージ密度、パフォーマンス、およびエラー訂正機能の条件に対応できなくなっています。これらの限界を克服するため、2009 年に各社共通の形式となる 4096 バイトの物理セクターフォーマットが定められました（「アドバンストフォーマット」、4K セクター）。このフォーマットを使用するには、影響を受けるあらゆるハードウェアおよびソフトウェアコンポーネントがサポートすることが前提とされます。これらのコンポーネントには次のものがあります。

- ・ HDD
- ・ ディスクコントローラー
- ・ オペレーティングシステム
- ・ ドライバ、アプリケーション
- ・ その他ツール。

メーカーによるこれらすべてのコンポーネントの移行には当然ながらある程度の時間がかかるため、暫定的に 512 バイトセクターのエミュレーションテクノロジーが導入されています。つまり、HDD の内部は 4K 物理セクターでありながら、外部から見ると 512 バイトセクターにエミュレートしたものとして扱われます。これにより、新しいハードディスクも既存の環境で使用することが可能になります。これらの HDD は、略して「512e」で呼ばれることもあります（「e」は「エミュレーション」を意味する）。

PRIMERGY サーバに搭載されるハードディスクでは、4K テクノロジーは 512e HDD の形でのみ実現されています。この場合、512 バイトセクターのエミュレーションは常に有効化され、4K ネイティブへのスイッチモードは存在しません。

HDD 内部は、物理セクターサイズ単位でアクセスがされます。物理セクターサイズより小さい書き込みアクセスがあった場合は、下図のように「リード・モディファイ・ライトサイクル」が発生します。



- ① ホストからの 512 バイト (512b) セクターの書き込み要求
- ② HDD が 4KB セクターを DRAM に読み込み
- ③ HDD が DRAM 内の要求された 512 バイトセクターを更新
- ④ HDD がメディアの 4KB セクターを上書き

パフォーマンス

リード・モディファイ・ライトサイクルにより HDD 上で余分な動作が生じるため、パフォーマンスが低下します。このため、アプリケーションとオペレーティングシステム側で最高のパフォーマンスを実現させるには、物理セクターサイズより小さい書き込みアクセスを防止する必要があります。

以下の場合、パフォーマンスが低下します。

・書き込みアクセスサイズが 4K の倍数にならない書き込みアクセスが発生した場合

今日の 4K の倍数のサイズで書き込みアクセスを行うアプリケーションとオペレーティングシステムでは問題になりません。

・書き込みアクセスサイズが 4K の倍数になりながら、物理セクターとファイルシステムの境界が一致していない場合

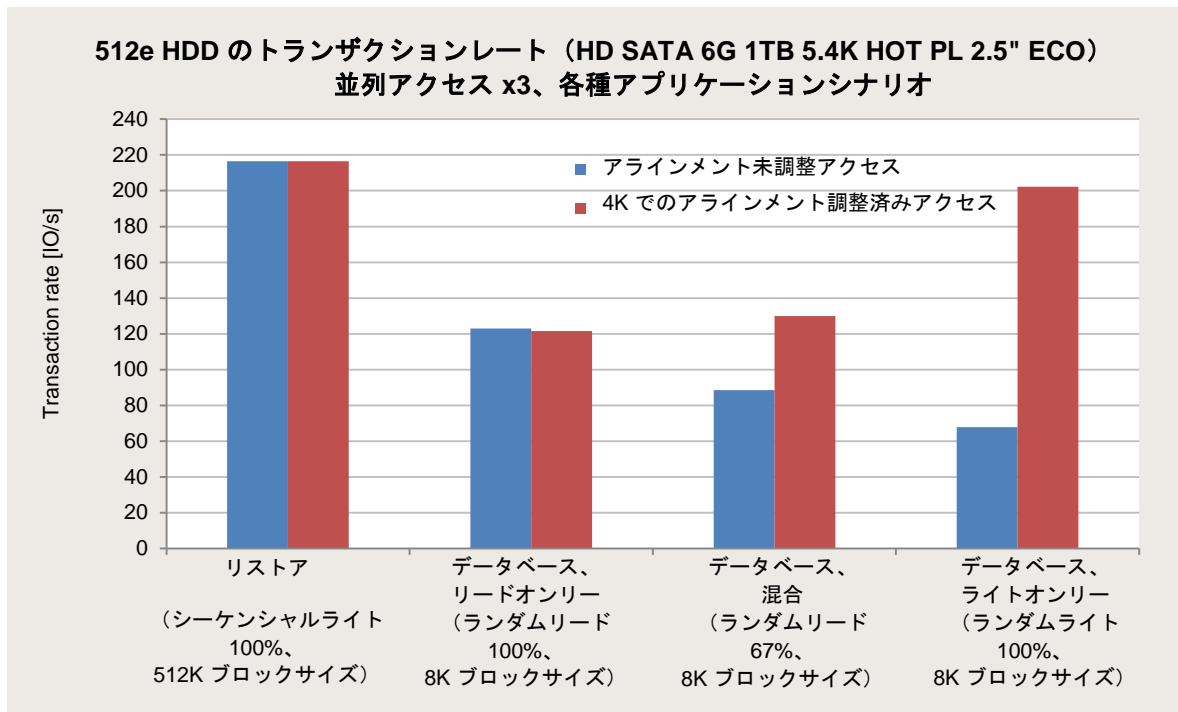
物理セクターとファイルシステムの境界を一致させる必要があります。これを、書き込みアクセスの調整（アライメントの調整）と呼びます。特に以下の 2 つの要件を満たす必要があります。

①パーティション作成時に、ファイルシステムの境界を物理 4k アライメントに調整する。

②パーティション作成後も、全てのソフトウェアコンポーネントからの書き込みアクセスを 4k アライメントに調整する。

物理 4k セクターを持つ HDD をさまざまな角度から考察した説明については、Microsoft の記事「[Advanced format \(4K\) disk compatibility update](#)」（アドバンストフォーマット (4k) ディスクの互換性に関する最新情報）」を参照してください。この記事では、物理 4k セクターに対するソフトウェアの適合性についても説明しています。※ 英語表記

次の図は、アラインメント調整済アクセスとアラインメント未調整アクセスの場合の 512e HDD のパフォーマンスの違いを示しています。



100% 読み取りの場合は、シーケンシャルアクセスとランダムアクセスともパフォーマンスに違いはありません。アラインメント調整済みアクセスが有利になるのは、書き込み時のランダムアクセスのみとなっています。図の 3 列目のグループの例（データベースに典型的な頻繁なアクセスプロファイルに相当）では、書き込みが 33% を占めています。このため、アラインメント調整済みアクセスでは、書き込み率が大きいほどその利点も大きくなります。100% ライトのランダム負荷プロファイルでは、整合アクセスによって不整合アクセスの実に 3 倍のパフォーマンス優位性が得られます。これは図の 4 列目のグループが示しています。

補足 :

一部のハードディスクでは、追加の内部メカニズムによってアラインメント未調整書き込みアクセスの影響を緩和しています。（例、Seagate 社の「[SmartAlign](#)」テクノロジー）

注意が必要なシステム構成

最適なパフォーマンスを得るために、4k セクターに対応した以下のディスクコントローラー、ドライバおよびオペレーティングシステムを適用する必要があります。

ディスクコントローラー			
タイプ	製品型名	製品名称	ファームウェアパッケージバージョン
SAS アレイコントローラ	PY-SR2C2 / PYBSR2C2	RAID Ctrl SAS 6G 5/6 512 MB (D2616)	≥ 12.12.0-0129
	PY-SDR08 / PYBSRD082	PY SAS RAID Mezz Card 6Gb	
	PY-SR2W0 / PYBSR2W0 / PYBSR2W0L	MegaRAID SAS 9280-8e	
	PY-SR2L2 / PYBSR3L2	RAID Ctrl SAS 6G 0/1 (D2607)	
	PY-SR3C3 / PYBSR3C3	RAID Ctrl SAS 6G 5/6 1GB (D3116)	
	PY-SR3C32 / PYBSR3C32	RAID Ctrl SAS 6G 5/6 1GB (D3116C)	
	PY-SR3C33 / PYBSR3C33	PY SAS RAID HDD Module (D2816)	
	PY-SR3C34 / PYBSR3C34	PY SAS RAID HDD Module (D2816C)	
	PY-SDR24 / PYBSRD241	PY SAS RAID HDD Module w/o cache (D2837)	
	PY-SR3D24A / PYBSRD241A	PY SAS RAID HDD Module w/o cache (D2837C)	
オンボード RAID	PY-SR14 / PYBSRD141	MegaRAID SAS 9285CV-8e	≥ 23.29.0-0019
	PY-SR14A / PYBSRD141A	MegaRAID SAS 9286CV-8e	
	PY-SR3PR / PYBSR3PR / PYBSR3PRL	PRAID EP400i	
	PY-SR3PR2 / PYBSR3PR2 / PYBSR3PRL2	PRAID EP420i	
非 RAID	PY-SR3FA / PYBSR3FA	PRAID CP400i	ファームウェアによる影響なし
	PY-SR3PE / PYBSR3PE / PYBSR3PEL	PRAID EP420e	
	オンボードソフトウェア RAID	Embedded MegaRAID	
	PY-SCD08 / PYBSCD082	SAS HBA Mezzanine Card (D3007)	
オペレーティングシステム	PY-SC2L2 / PYBSC2L2	PSAS CP200i	ファームウェアによる影響なし
	PY-SC2Z0 / PYBSC2Z0 / PYBSC2Z0L	LSI SAS9200-8e	
	PY-SC3FE / PYBSC3FE / PYBSC3FEL	PSAS CP400e	

オペレーティングシステム	ドライババージョン
Windows	SAS アレイコントローラ : MegaRAID sas ≥ 6.505.5 オンボードソフトウェア RAID : MegaSR ≥ 15.02.2013.0425 SAS コントローラ (非 RAID モード) : すべて
Linux	SAS アレイコントローラおよび SAS コントローラ (非 RAID モード) : ネイティブサポート オンボードソフトウェア RAID : MegaSR ≥ 15.02.2013.0425

オペレーティングシステム	
名前	追加の要件
Windows Server 2008	KB2553708
Windows Server 2008 R2	KB982018 または Service Pack 1
Windows Server 2012	-

RHEL	≥バージョン 5
SLES	≥バージョン 11 SP2

国または販売地域によっては、一部のコンポーネントが利用できない場合があります。

コメント :

- リリースされているコントローラーとハードディスクの組み合わせについては、システム構成図をご覧ください。
- オペレーティングシステムのインストール時に最適なパフォーマンスを確保するには、上記のコンポーネントを適切な段階で適用する必要があります。
- 通常のオペレーティングシステム環境以外（例：イメージングツールを使用したパーティション作成またはハードディスク全体のクローン作成など）で、最適なパフォーマンスを実現するには、これらツールも 4k セクターをサポートする必要があります。

著作権・商標権・その他の知的財産権について

コンテンツ（文書・画像・音声等）は、著作権・商標権・その他の知的財産権で保護されています。本コンテンツは、個人的に使用する範囲でプリントアウトまたはダウンロードできます。ただし、これ以外の利用（ご自分のページへの再利用や他のサーバへのアップロード等）については、当社または権利者の許諾が必要となります。

保証の制限

本コンテンツについて、当社は、その正確性、商品性、ご利用目的への適合性等に関する保証するものではなく、そのご利用により生じた損害について、当社は法律上のいかなる責任も負いかねます。本コンテンツは、予告なく変更・廃止されることがあります。