

# ホワイトペーパー

## Fujitsu PRIMERGY サーバ

### ソリッドステートドライブ - FAQ

フラッシュメモリをベースとするソリッドステートドライブ（SSD）は、ここ数年、従来型ハードディスクの代替として使用されてきました。SSD を導入する際には、ストレージメディアの信頼性、システムの応答速度、システム効率性の要件を基準に検討します。もちろん、コストも重要な決定要因です。このドキュメントは、SSD に関してよく寄せられる質問とその回答（FAQ）をまとめたものです。従来型 SSD、PCIe-SSD と DOM（ディスクオンモジュール）の技術面について、その寿命に重点を置いて解説するとともに、PRIMERGY サーバファミリーでの SSD の最適な使用方法についても説明します。

#### バージョン

1.1

2014-03-31



## 目次

ドキュメントの履歴 .....	3
全般 .....	4
内蔵 SSD に関する FAQ .....	4
SSD と HDD の基本的な違いを教えてください。 .....	4
フラッシュメモリにどのようにアクセスしますか。 .....	4
ライトアンプリフィケーションとは何ですか。 .....	5
ライトアンプリフィケーションは、パフォーマンスや書き込み耐久性にどのような影響を及ぼしますか。 .....	6
SSD はデフラグできますか。 .....	6
コマンド TRIM にはどのような機能がありますか。 .....	6
コマンド TRIM が機能するのはどのような場合ですか。 .....	7
TRIM が機能しないと、どのような影響が生じると考えられますか。 .....	7
オーバープロビジョニングの意味を教えてください。 .....	7
ドライブ クラスについての FAQ .....	9
富士通が提供するドライブクラスについて教えてください。 .....	9
PRIMERGY サーバ向けの SSD はどれですか。 .....	10
DOM とは。 .....	11
PCIe-SSD が従来型 SSD (SATA/SAS) と異なるのはどのような点ですか。 .....	11
SSD によって価格が大きく違うのはなぜですか。 .....	12
フェールセーフと書き込み耐久性に関する FAQ .....	13
セキュリティは、SSD よりも HDD の方が優れているのですか。 .....	13
回復不能な読み取りエラーとは何ですか。 .....	13
KPI の MTBF と AFR は何を意味しているのですか。 .....	13
SSD と HDD のフェールセーフにはどのような差異がありますか。 .....	14
書き込み耐久性は何を意味しているのですか。 .....	15
DWPD 値と、1 日あたりの実際の書き込み負荷との関係を教えてください。 .....	16
現在使用している SSD を、書き込み耐久性が同等以上の SSD と交換したいと考えています。どうすれば、適切な SSD を選択できますか。 .....	16
RAID 構成の選択は、SSD の書き込み耐久性に影響しますか。 .....	17
書き込み耐久性の期間が終了したらどうすればよいですか。 .....	17
SSD へのアクセスがない場合、データが保持される期間はどの程度ですか。 .....	17
接続されていない SSD はどの程度の期間データを保持できますか。 .....	18
インターフェースに関する FAQ .....	19
SATA-SSD ではなく SAS-SSD を使用した方がよいのはどのような場合ですか。 .....	19
RAID に関する FAQ .....	20
SSD を使用するには RAID が必要ですか。 .....	20
HDD から SSD に変換する場合、RAID 構成について注意すべきことはありますか。 .....	20
SSD または PCIe-SSD を使用する場合の推奨 RAID レベルを教えてください。 .....	20
HW-RAID と SW-RAID .....	20
SSD を使用する場合の推奨コントローラーを教えてください。 .....	21
キャッシュ付きの RAID コントローラーが必要なのはどのような場合ですか。 .....	21
SSD を使用する場合、ホットスペアは必要ですか。 .....	21
さまざまな用途における SSD と HDD に関する FAQ .....	22
HDD と SSD のパフォーマンスの違いはどの程度ですか。 .....	22
このホワイトペーパーでは私の疑問に対する答えが見つかりません。なぜですか。 .....	24

関連資料.....	25
お問い合わせ先.....	25

## ドキュメントの履歴

### バージョン 1.0

初版

### バージョン 1.0a

- マイナー修正
- 「[キャッシュ付きの RAID コントローラーが必要なのはどのような場合ですか](#)」という質問について、RAID 1 に関する例外は適用されなくなったため削除しました。

### バージョン 1.1

- ドライブクラスの更新
- 耐久性クラスの更新
- 新しい SSD モデルの追加
- DOM の追加
- パフォーマンス値の更新

## 全般

通常のドキュメントと同様に、このホワイトペーパーでも、容量の仕様には SI 規格に基づく 10 進接頭辞（1 kB =  $10^3$  バイト、1 MB =  $10^6$  バイト、...、1 EB =  $10^{18}$  バイト）を使用しています。

ブロックサイズとスループットの仕様には、慣行に従い 2 進接頭辞を使用していますが、10 進接頭辞と区別できるように、IEC 規格に基づく 2 進接頭辞で表記しています（1 KiB =  $2^{10}$  バイト、1 MiB =  $2^{20}$  バイト、...、1 TiB =  $10^{40}$  バイト）

単位が大きくなるほど、2 進接頭辞と 10 進接頭辞の数値の差異も大きくなる点に留意してください。

1 KiB = 1.024 kB = 1 kB + 2.4%

1 TiB ≈ 1.1 TB = 1 TB + 10%

## 内蔵 SSD に関する FAQ

### SSD と HDD の基本的な違いを教えてください。

SSD は従来型のハードディスクと違って可動部分がなく、フラッシュメモリをベースとしています。従来型 SSD の外形とインターフェースは、従来型ハードディスクと同様です。一方、PCIe-SSD は、その名のとおりに、PCIe インターフェースを通じて直接作動します。SSD は、従来型ハードディスクと比べて、高速アクセス、機械的な堅牢性、低ノイズ、低消費電力の面で優れていますが、現在はまだ容量が小さく高価です。

SSD はフラッシュメモリセルに情報を保存しますが、これにはさまざまなタイプがあります。

- シングルレベルセル（SLC）は 1 セルあたり 1 ビット保存します。50 nm テクノロジーの SLC は、1 セルあたり約 100,000 回の書き込み操作が可能です。このタイプのセルを使用した SSD は一般的に小容量ですが、書き込み性能と書き込み耐久性に優れています。
- マルチレベルセル（MLC）は 1 セルあたり 2 ビット保存します。書き込み可能回数はサイズによって異なります。25 nm テクノロジーの MLC の場合、1 セルあたりの書き込み可能回数は約 3,000 回です。MLC ベースの SSD は、SLC ベースの SSD に比べて、大容量ですが書き込み性能と書き込み耐久性は劣ります。
- トリプルレベルセル（TLC）は 1 セルあたり 3 ビット保存します。25 nm テクノロジーの TLC の場合、1 セルあたりの書き込み可能回数は約 1,000 回です。容量は大きくなりますが、書き込み性能と書き込み耐久性は低下します。

SSD はアクセス速度が HDD よりも優れているので、特にサーバへの使用に適しています。実際には、SSD の書き込み回数には限界がありますが、SSD 内部のインテリジェントコントローラーによって、書き込み操作がすべてのフラッシュメモリに均一に分散されるので、消耗も均一に進みます。また、SSD モデルの多くは、内部ストレージの容量が公称値より大きくなっています。このような理由から、サーバに適しているだけでなく、HDD に真に匹敵する SSD モデルもあります。

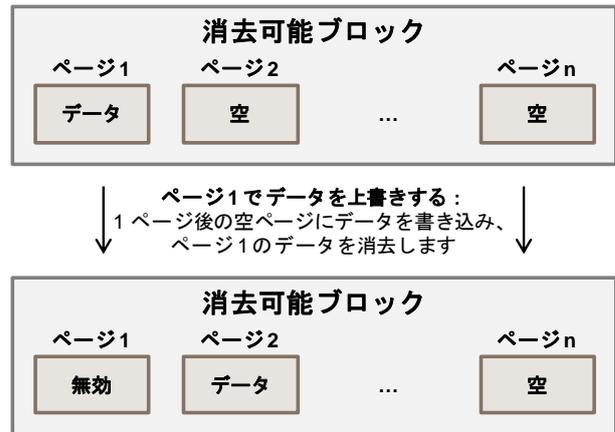
### フラッシュメモリにどのようにアクセスしますか。

SSD の読み取りおよび書き込み操作の最小単位は、1 ページで各ページのサイズは 4 Kib です。必要データ量に関係なく、必ず 4 KiB 以上のデータが読み取られます。書き込みジョブも、ブロックサイズに関係なく、4 KiB 単位で書き込み操作が実行されます。

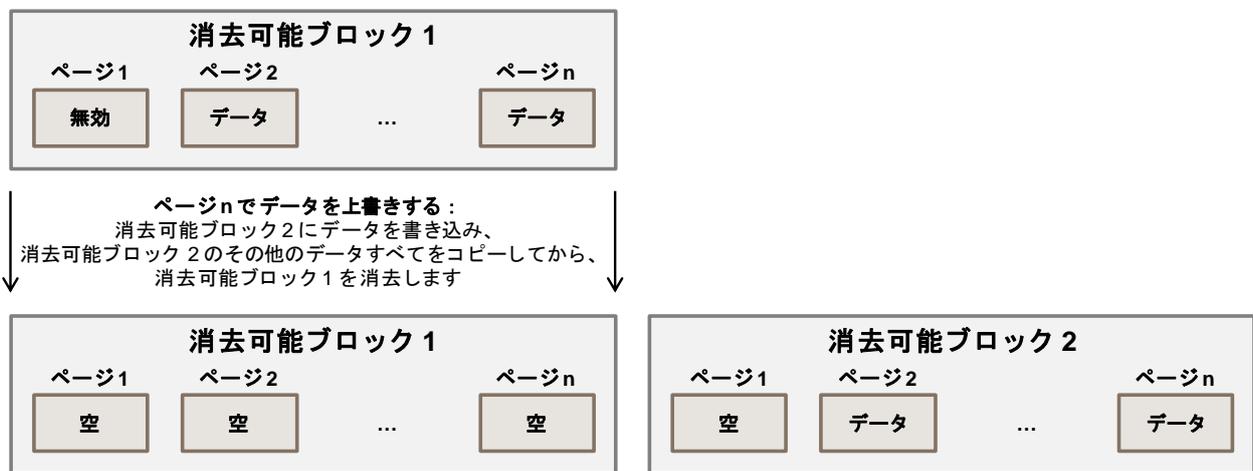
フラッシュメモリセルが空の場合に限り、書き込みが可能です。そのため、再利用するには、事前にフラッシュメモリセルを消去する必要があります。この動作はプログラム/消去サイクルと呼ばれています。各フラッシュメモリセルに実行できるプログラム/消去サイクル数には限界があるので、物理メモリに割り当てられているファイルが頻繁に更新されると、保存先のフラッシュセルが他のセルよりも早く消耗します。つまり、一部のメモリ領域が更新不能になっているのに、他のメモリ領域はまったく未使用であるという状態が生じる可能性があります。このような状態を回避するため、SSD のコントローラーはウェアレベリングと呼ばれる方法で、書き込みブロックを複数のメモリセルに分散させて、SSD の消耗を均一化します。このような場合、内部予備用のメモリセルも使用されます（[オーバープロビジョニング](#)を参照）。ウェアレベ

リングによって消耗は均一になりますが、その代償として、書き込み操作が増大します（[ライトアンプリフィケーション](#)を参照）。この現象を外部から制御することはできません。

SSD のウェアレベリングは、物理アドレスに論理ブロックアドレス（LBA）をマッピングすることを前提としています。データの更新では、元の物理アドレスは無効としてマークされ、更新には新しい物理アドレスが使用されます。アドレスが無効になっているページは、SSD によって事前に消去されないと再利用できません。これについて次に詳しく説明します。まず、SSD 内でページが統合され、**消去可能ブロック**が形成されます。ページのコンテンツが変更されると、古いページが無効としてマークされ、新しいコンテンツは消去可能ブロックのうち最も使用回数の少ない空きページに書き込まれます。



その消去可能ブロック内に使用可能な空きページがない場合は、空いている消去可能ブロックに無効ページがコピーされ、元の消去可能ブロックは消去されます。



このような方法で、無効になったページは書き込みジョブに再び利用できるようになります。この動作は**ガベージコレクション**と呼ばれています。この動作によって、ページ更新によるフラッシュメモリの消耗が均一化されます（**ダイナミックウェアレベリング**）。また、更新されない消去可能ブロックのフラッシュメモリセルも、均等に消耗するように一定の間隔でデータが自動的に移動されます（**スタティックウェアレベリング**）。

### ライトアンプリフィケーションとは何ですか。

SSD への実際の書き込み操作回数は、書き込みジョブのサイズが示す書き込み操作回数よりも明らかに大きくなる可能性があります。特に、ブロックサイズが小さいランダムアクセスの場合は、実際の操作回数が数倍になることもあります。この効果をライトアンプリフィケーション（または書き込み増幅）といいます。ライトアンプリフィケーションは、論理的には次のように計算されます。

$$\text{ライトアンプリフィケーション} = \frac{\text{SSD に物理的に書き込まれるデータボリューム}}{\text{論理的な書き込みジョブのデータボリューム}}$$

しかし、実際には、ライトアンプリフィケーションの程度はさまざまな要因の影響を受けるため、この数値は推定にすぎません。

## ライトアンプリフィケーションは、パフォーマンスや書き込み耐久性にどのような影響を及ぼしますか。

ライトアンプリフィケーションが高いと、SSD の書き込みジョブのパフォーマンスや書き込み耐久性にマイナスの影響を及ぼします。データブロックの小さいランダム書き込みでは特にライトアンプリフィケーションがかなり高くなり、その結果、書き込み操作回数が増えて、フラッシュセルの消費が速くなります。内蔵 SSD コントローラーのオーバープロビジョニングやインテリジェントデータ管理の機能はこの現象への対策として役立ちます。ここでは特に、データを均等に分散してフラッシュセルの消費を均一化するウェアレベリングについて説明します。ウェアレベリングは、エンタープライズクラス（「ドライブクラス」を参照）の従来型ハードディスクの一般的な寿命に匹敵する程度まで SSD の寿命を延ばす効果があります。

SSD メーカーの書き込み耐久性の仕様およびその KPI (TBW、PBW、DWPD) には、ランダムアクセスにおける標準書き込み負荷の典型的なライトアンプリフィケーションが考慮されています。つまり、書き込み負荷が異なれば、その結果として書き込み耐久性も異なります。負荷の差異が大きいほど耐久性の差も大きくなります。

## SSD はデフラグできますか。

SSD はデフラグ可能ですが、絶対に回避すべきです。

デフラグは HDD を想定したものであり、HDD 以外には実行する意味がありません。HDD のファイルは、最初は連続的に使用され、断片化しませんが、時間の経過とともに断片化が生じ、それによって徐々にパフォーマンスが低下します。当初のパフォーマンスレベルが回復されるので、デフラグ操作には意味があるといえます。

SSD の場合は、データが連続的であっても断片化していてもパフォーマンスに影響しません。このため、デフラグを実行してもメリットはありません。デフラグに不必要な時間をかけることになるうえ、デフラグを行うと、ライトアンプリフィケーションが高くなるため、SSD の書き込み耐久性にも悪影響が生じます。そのため、Windows Server 2012 がドライブのデフラグによって SSD に悪影響があると認識した場合、デフラグは実行されません。これは、PCIe-SSD が使用されている場合には認識されますが、RAID コントローラーを使用する従来型の SSD では認識されません。したがって、従来型の SSD を使用している場合、システム管理者は手動のデフラグもスケジュールタスクとしてのデフラグも実行されないように手段を講じる必要があります。

## コマンド TRIM にはどのような機能がありますか。

TRIM は、論理レベルで消去され、物理的にはデータ メディア上に残っているデータの消去を目的としたコマンドです。

背景：

ファイルの消去や、Windows の Quick フォーマットなどのハイレベルフォーマットを実行すると、前に保存していたデータは参照できなくなります。しかし、データ自体はデータメディアから消去されているとは限りません。HDD の場合はこれが問題となることはなく、このような複雑な消去の仕組みは不要です。しかし、SSD ではこれが問題となります。この場合、SSD は実際に使用可能なメモリ領域よりも少なく認識します。論理的に消去されても物理的にデータが残っているメモリ領域を、空のブロックとしてダイナミックウェアレベリングに使用することはできません。また、空のブロックとは異なり、これらのメモリ領域はスタティックウェアレベリングの対象となるため、ライトアンプリフィケーションも増大します。このようなメモリ領域をサーバが再利用するには、フラッシュメモリセルの消去が不可欠ですが、これは新しい書き込みジョブの直前に実行されることとなります。そのため、このようなフラッシュメモリセルへの書き込み操作では応答に時間がかかります。

これを回避するため、新しいオペレーティング システム バージョンでは、ファイル消去時の物理的なデータ消去や、SSD が論理ドライブの一部として認識された場合のハイレベルフォーマットに、コマンド TRIM が使用されます。

## コマンド TRIM が機能するのはどのような場合ですか。

### 従来型 SSD :

RAID コントローラーは TRIM コマンドをサポートしていないので、従来型 SSD ではこの機能は使用できません。

### PCIe-SSD :

TRIM コマンドは通常、Windows と Linux ではサポートされています。

Windows :

- RAID 5 を除く RAID ボリューム（ミラーリング、スパンニング、またはストライピングを利用）
- シンプルボリューム
- 複数のデータメディアにおける上記の RAID レベルの各組合せ（PCIe-SSD を 1 台以上含む）
- RAID アレイ内の PCIe-SSD
- PCIe-SSD の複数のパーティション
- NTFS および FAT32 ファイルシステム
- マウントポイントがあるボリューム
- 圧縮されたボリューム
- さまざまなクラスタサイズ、パッケージサイズ、およびセクターサイズ
- いわゆる「拡張された」ボリュームおよび「縮小された」ボリューム

Linux :

- 「Discard」をサポートしているすべてのオペレーティング システムで、
- 「Discard」が有効な場合
- ドライブを作成できる Linux コマンドのすべてが「Discard」を実行するわけではありません（該当する Linux ディストリビューションのマニュアルを参照）。

## TRIM が機能しないと、どのような影響が生じると考えられますか。

[ライトアンプリフィケーション](#)が増大し、すでに消去されているメモリ領域にアクセスする際のパフォーマンスが低下します。

## オーバープロビジョニングの意味を教えてください。

SSD には通常、内部予備用のフラッシュメモリセルがあります。これは書き込み操作時にバッファとして使用されるほか、書き込み可能なメモリセルがない場合に交換用のメモリセルとして使用されます。これは、SSD の書き込み性能と[書き込み耐久性](#)にプラスの影響を与えます。通常、この予備の大きさは SSD 容量のパーセンテージで指定されます。

$$\text{SSD 容量のパーセンテージで示されるオーバープロビジョニング} = \frac{\text{物理的容量} - \text{論理的容量}}{\text{論理的容量}} \times 100$$

オーバープロビジョニング領域のサイズは SSD モデルによってさまざまです。一般的に SSD メーカーは明記していませんが、[DWPD および PBW（または TBW）](#)についてのメーカー情報で考慮されています。

一部の SSD メーカーの SSD、特にコンシューマー製品の SSD には、オーバープロビジョニング領域がまったくないものもあります。この場合、顧客が SSD 容量の 70 % だけ使用するなど、独自のオーバープロビジョニングを設定することが想定されています。

## ドライブクラスについての FAQ

### 富士通が提供するドライブクラスについて教えてください。

富士通は、考えられるあらゆるシナリオ（業務時間のみ稼動すればよい小規模部門用のサーバから高可用性のデータベースサーバまで）に対応するため、PRIMERGY サーバ向けに、パフォーマンスや可用性の特徴が異なるさまざまなストレージメディア（HDD および SSD）を提供しています。豊富な種類の中から適切な製品を簡単に選択できるように、富士通は次の 5 種類のドライブクラスを紹介しています。

ドライブクラス	説明/適合性	タイプ	容量 <sup>1)</sup>
エコノミック (ECO)	低価格。ただし、提供されるパフォーマンスと信頼性は、エントリーレベルの用途のみを対象としています。重要性が低く、低 I/O トラフィック、中速度で問題のない領域で使用してください。高負荷では信頼性が損なわれる可能性があります。ECO ドライブは、回転速度 5400/7200 rpm で作動し、SATA インターフェースを搭載しています。	2.5 インチ SATA 5400 HDD	1 TB
		3.5 インチ SATA 7200 HDD	250 GB~500 GB
ビジネスクリティカル (BC)	1 GB あたり最小コストで最大の容量を提供。高性能およびそれに対応する信頼性を実現できるように設計されています。サーバの実装によっては、BC ドライブと一緒に SAS または SATA インターフェースが組み込まれることもあります。回転速度は 7200 rpm です。最高レベルの I/O スループットが必要とされる環境では、代替としてエンタープライズクラスの HDD または SSD を使用する必要があります。	2.5 インチ SATA 7200 HDD	250 GB~1 TB
		2.5 インチ SAS 7200 HDD	500 GB~1 TB
		3.5 インチ SATA 7200 HDD	500 GB~4 TB
		3.5 インチ SAS 7200 HDD	1 TB~4 TB
エンタープライズ (EP) <sup>2)</sup>	最大のパフォーマンスと信頼性を提供。高いスループットと少ない遅延を実現します。	2.5 インチ SAS 10000 HDD	300 GB~1.2 TB
		2.5 インチ SAS 15000 HDD	146 GB~300 GB
		2.5 インチ SSD	100 GB~1.6 TB
		3.5 インチ SAS 15000 HDD	300 GB~600 GB
		PCIe-SSD	365 GB~1.2 TB
		DOM	64 GB

<sup>1)</sup> 日付 : 2014-04-01

<sup>2)</sup> 以前 : エンタープライズ (EP) 、エンタープライズメインストリーム (EP MAIN) およびエンタープライズパフォーマンス (EP PERF)

**PRIMERGY サーバ向けの SSD はどれですか。**

名前	インターフェース	容量 [GB]	書き込み耐久性 [PBW]	DWPD (切り捨て)	PBW に達するまでの MiB/s	回復不能な読み取りエラー (ビットエラーレート)	MTBF [百万時間]	形状
PCIe-SSD 1.2TB MLC	PCIe Gen2 x4	1200	17	7	103	1/10 <sup>20</sup>	1.5	LP <sup>1)</sup>
PCIe-SSD 785GB MLC	PCIe Gen2 x4	785	11	7	67	1/10 <sup>20</sup>	2	LP <sup>1)</sup>
PCIe-SSD 640GB MLC <sup>2)</sup>	PCIe Gen1 x4	640	10	8	60	1/10 <sup>20</sup>	1.24	LP <sup>1)</sup>
PCIe-SSD 365GB MLC	PCIe Gen2 x4	365	4	6	24	1/10 <sup>20</sup>	2	LP <sup>1)</sup>
PCIe-SSD 320GB MLC <sup>2)</sup>	PCIe Gen1 x4	320	4	6	24	1/10 <sup>20</sup>	1.24	LP <sup>1)</sup>
SSD SAS 12G 1.6TB Main 2.5" H-P EP	SAS 12G	1600	29.2	10	176	1/10 <sup>17</sup>	2	2.5"
SSD SAS 12G 800GB Main 2.5" H-P EP	SAS 12G	800	14.6	10	88	1/10 <sup>17</sup>	2	2.5"
SSD SAS 12G 400GB Main 2.5" H-P EP	SAS 12G	400	7.3	10	44	1/10 <sup>17</sup>	2	2.5"
SSD SAS 12G 200GB Main 2.5" H-P EP	SAS 12G	200	3.65	10	22	1/10 <sup>17</sup>	2	2.5"
SSD SAS 6G 400GB SLC HOT P 2.5" EP PERF <sup>2)</sup>	SAS 6G	400	35	47	212	1/10 <sup>16</sup>	2	2.5"
SSD SAS 6G 400GB MLC HOT PL 2.5" EP PERF <sup>2)</sup>	SAS 6G	400	7.5	10	45	1/10 <sup>16</sup>	2	2.5"
SSD SAS 6G 200GB SLC HOT P 2.5" EP PERF <sup>2)</sup>	SAS 6G	200	18	49	109	1/10 <sup>16</sup>	2	2.5"
SSD SAS 6G 200GB MLC HOT PL 2.5" EP PERF <sup>2)</sup>	SAS 6G	200	3.75	10	23	1/10 <sup>16</sup>	2	2.5"
SSD SAS 6G 100GB SLC HOT P 2.5" EP PERF <sup>2)</sup>	SAS 6G	100	9	49	54	1/10 <sup>16</sup>	2	2.5"
SSD SAS 6G 100GB MLC HOT PL 2.5" EP PERF <sup>2)</sup>	SAS 6G	100	1.875	10	11	1/10 <sup>16</sup>	2	2.5"
SSD SATA 6G 800GB Main 2.5" H-P EP	SATA 6G	800	14.6	10	88	1/10 <sup>17</sup>	2	2.5"
SSD SATA 6G 400GB Main 2.5" H-P EP	SATA 6G	400	7.3	10	44	1/10 <sup>17</sup>	2	2.5"
SSD SATA 6G 400GB MLC HOT P 2.5" EP MAIN <sup>2)</sup>	SATA 6G	400	7.5	10	45	1/10 <sup>16</sup>	2	2.5"
SSD SATA 6G 200GB Main 2.5" H-P EP	SATA 6G	200	3.65	10	22	1/10 <sup>17</sup>	2	2.5"
SSD SATA 6G 200GB MLC HOT P 2.5" EP MAIN <sup>2)</sup>	SATA 6G	200	3.75	10	23	1/10 <sup>16</sup>	2	2.5"
SSD SATA 6G 100GB Main 2.5" H-P EP	SATA 6G	100	1.825	10	11	1/10 <sup>17</sup>	2	2.5"
SSD SATA 6G 100GB MLC HOT P 2.5" EP MAIN <sup>2)</sup>	SATA 6G	100	1.875	10	11	1/10 <sup>16</sup>	2	2.5"
SSD SATA 3G 64GB SLC HOT P 2.5" EP MAIN <sup>2)</sup>	SATA 3G	64	2	17	12	1/10 <sup>15</sup>	2	2.5"
SSD SATA 3G 32GB SLC HOT P 2.5" EP MAIN <sup>2)</sup>	SATA 3G	32	1	17	6	1/10 <sup>15</sup>	2	2.5"
DOM SATA 3G 64GB Main N H-P	SATA 3G	64	1.6425	14	9	n/a	3	DOM <sup>3)</sup>

<sup>1)</sup> LP : ロープロファイル

<sup>2)</sup> EOL : 製造終了

<sup>3)</sup> DOM : ディスクオンモジュール

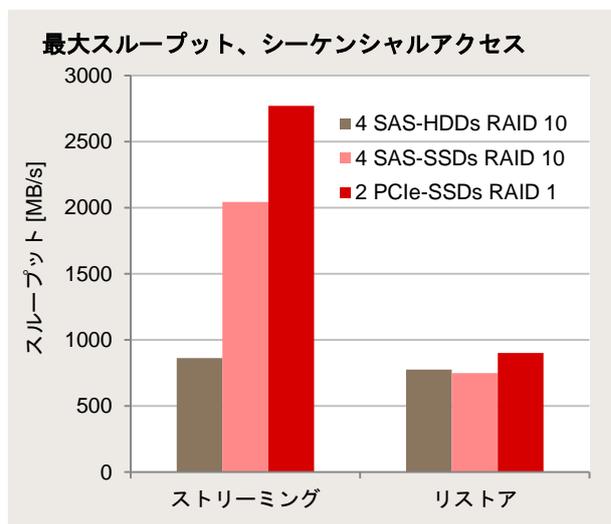
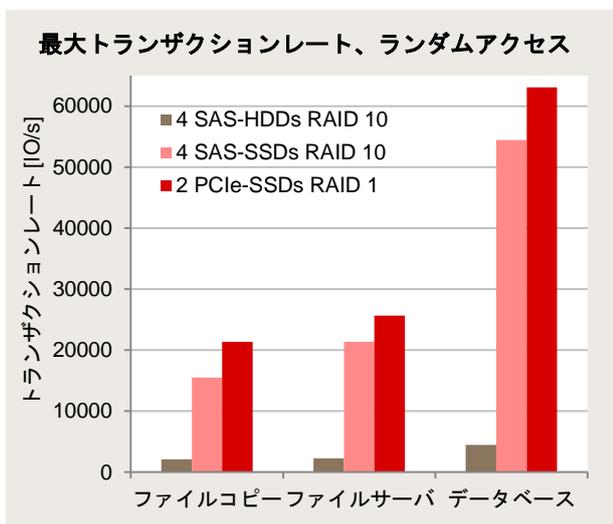
## DOM とは。

DOM は「Disk on module : ディスクオンモジュール」の略です。省スペースと省電力の効果が高いフラッシュメモリで、特にサーバのブートドライブとして使用されます。このメモリテクノロジーは SSD と同じです。Fujitsu の PRIMERGY サーバのラインアップには、SATA 3G インターフェースの DOM が提供されていて、システムボードの SATA ポートに直接挿入できます。

## PCIe-SSD が従来型 SSD (SATA/SAS) と異なるのはどのような点ですか。

従来型 SSD はホストバスアダプタ (通常 RAID コントローラー) を通じて動作し、SATA または SAS インターフェースが搭載されています。システムボードのチップセットに対する RAID コントローラーのインターフェースは、通常 PCIe か、または統合型オンボードコントローラーの場合はシステムボードの内部バスインターフェースです。一方、PCIe-SSD は PCIe インターフェースのみを介して直接動作します。

個々の PCIe-SSD を 1 つの従来型 SSD と比較すると、パフォーマンスは前者の方がはるかに優れています。したがって、PCIe-SSD は従来型 SSD を使用した RAID 構成の代替として十分に使用できます。たとえば、1 つの RAID 内にある 2 つの PCIe-SSD は、RAID 10 内の 4 つの SAS-SSD よりもパフォーマンスが上回る可能性があります。詳細については、『[RAID Controller Performance](#)』と『[Performance Report PCIe-SSDs ioDrive®2](#)』を参照してください。



## SSD によって価格が大きく違うのはなぜですか。

HDD と同様、SSD の価格はその容量、インターフェース、パフォーマンスによって左右されます。もう 1 つ、SSD 固有の基準である [書き込み耐久性](#) も重要です。SSD は耐久性に関して大きく 3 つのカテゴリーに分けることができます。対応できる要件はカテゴリーによってまったく異なります。

耐久性クラス	DWPD	説明/適合性
バリュー耐久性 (読み取り集中)	< 約 5 通常、< 0.3	主に MLC フラッシュメモリを使用した低価格カテゴリーの SSD。通常、平均 3 MiB/s 未満の書き込み負荷を 5 年以上維持可能です。これは、低い書き込み強度が特徴の負荷状況に適しています。 例：システムドライブ、ストリーミングサービス
主流の耐久性	約 5~約 15 通常、約 10	主に MLC フラッシュメモリを使用した、中間価格カテゴリーの SSD と、優れたパフォーマンスのために高価格の PCIe-SSD、そして DOM。 通常、平均 2 桁の MiB/s 範囲にある書き込み負荷を 5 年以上維持可能です。これは、中程度の書き込み強度が特徴の負荷状況に適しています。 例 (SSD) : ファイルサーバ、Web サーバ 例 (PCIe-SSD) : 仮想サーバ、データベース、メールサーバ 例 (DOM) : システムドライブ
高耐久性 (書き込み集中)	> 約 15 多くの場合、約 50	主に SLC フラッシュメモリを使用した高価格カテゴリーの SSD。通常、平均 2~3 桁の MiB/s 範囲にある書き込み負荷を 5 年以上維持可能です。これは、高い書き込み強度が特徴の負荷状況に適しています。 例：仮想サーバ、データベース、メールサーバ

一般的に「エンタープライズ SSD」という場合、その SSD に指定されている耐久性クラスは明確ではありません。メーカーの多くは主流クラスや高耐久性クラスの SSD にこの名称を使用していますが、バリュー耐久性クラスの SSD や SSD 全般にこの名称が使用されることもあります。

## フェールセーフと書き込み耐久性に関する FAQ

### セキュリティは、SSD よりも HDD の方が優れているのですか。

サーバ環境への SSD の導入に反対する議論では、SSD は HDD よりもセキュリティが弱いという主張がよく聞かれます。全体的に見ただけでは、この主張の是非を判断することはできません。セキュリティの個々の側面を個別に検討する必要があります。

- フェールセーフ (回復不能な読み取りエラー、MTBF、AFR) :  
3つの PCIe-SSD モデルを除き、SSD と HDD は差異はありません。
- [書き込み耐久性](#) :  
SSD 固有の基準であり、PRIMERGY サーバ向けの SSD モデルはすべての要件を満たしています。
- アーカイブ  
SSD は長期アーカイブ用のメディアではありません。短期のアーカイブには使用できます。

ここでは質問に回答する形で、この問題について詳しく説明します。

### 回復不能な読み取りエラーとは何ですか。

いわゆる回復不能な読み取りエラーがデータメディアの障害の原因になることがよくあります。この現象の発生頻度は、HDD や SSD の品質に関する重要な指標です。

回復不能な読み取りエラーはストレージメディアの障害によって発生します。このような致命的なデータ損失は、RAID アレイにデータを重複保存することで回避できます。ただし、回復不能な読み取りエラーは RAID アレイの再構築に重大な影響を及ぼすこともあります。特に、データメディアあたりの容量が大きいと再構築に要する時間が長くなります。回復不能な読み取りエラーは RAID 5 アレイの再構築に致命的な影響を及ぼすため、RAID 6 や RAID 10 のソリューションが普及しつつあります。RAID 6 ソリューションは、RAID 5 ソリューションより価格はわずかに高くなり、性能もそれほど高くはありません。一方、RAID 10 ソリューションは、高性能で再構築時間も短縮されますが、かなり高価格になります。

RAID コントローラには「パトロールリード」という機能があります。これは、RAID アレイの不良ブロックを事前に追跡することにより、不良ブロックに最初にアプリケーションがアクセスしないようにする機能です。このような不良ブロックは不良としてマークされ、データストレージプールから除外されます。これはローレベルフォーマットでも同様です。これによって、回復不能な読み取りエラーの発生率を大幅に軽減できます。

### KPI の MTBF と AFR は何を意味しているのですか。

ストレージメディア (HDD または SDD) の寿命の推定には、一般的に MTBF 値が使用されます。MTBF (Mean Time Between Failures : 平均故障間隔) は、故障から次の故障までの平均稼働期間を時間単位で表します。これはあくまで統計上の数値であり、特定の状況でその MTBF 値の間故障が起きないことを保証するものではありません。

場合によっては、AFR (Annualized Failure Rate : 年間故障率) 指標が使用されることもあります。これは MTBF から算出される値で、

$$AFR=100 \times \frac{24 \times 365}{MTBF}$$

デバイス (ここでは HDD または SDD) が 1 年以内に故障するパーセンテージです。

反対に、

$$1 - AFR$$

はデバイスが 1 年間故障せずに作動する推定パーセンテージです。

例 : MTBF = 200 万時間の場合 : AFR = 0.44 % で、1 - AFR = 99.56 % です。

## SSD と HDD のフェールセーフにはどのような差異がありますか。

容量、パフォーマンス、価格だけでなく、フェールセーフも SSD を購入するかどうかの重要な判断基準です。エンタープライズ HDD と SSD の故障率はほぼ同じです。このため、RAID システムの構成に使用する RAID レベルを選択する際には、同じセキュリティ基準に基づいて判断する必要があります。

名前	MTBF [百万時間]
PCIe-SSD 1.2TB MLC	1.5
PCIe-SSD 785GB MLC	2
PCIe-SSD 640GB MLC *)	1.24
PCIe-SSD 365GB MLC	2
PCIe-SSD 320GB MLC *)	1.24
SSD SAS 12G 1.6TB Main 2.5" H-P EP	2
SSD SAS 12G 800GB Main 2.5" H-P EP	2
SSD SAS 12G 400GB Main 2.5" H-P EP	2
SSD SAS 12G 200GB Main 2.5" H-P EP	2
SSD SAS 6G 400GB SLC HOT P 2.5" EP PERF *)	2
SSD SAS 6G 400GB MLC HOT PL 2.5" EP PERF *)	2
SSD SAS 6G 200GB SLC HOT P 2.5" EP PERF *)	2
SSD SAS 6G 200GB MLC HOT PL 2.5" EP PERF *)	2
SSD SAS 6G 100GB SLC HOT P 2.5" EP PERF *)	2
SSD SAS 6G 100GB MLC HOT PL 2.5" EP PERF *)	2
SSD SATA 6G 800GB Main 2.5" H-P EP	2
SSD SATA 6G 400GB Main 2.5" H-P EP	2
SSD SATA 6G 400GB MLC HOT P 2.5" EP MAIN *)	2
SSD SATA 6G 200GB Main 2.5" H-P EP	2
SSD SATA 6G 200GB MLC HOT P 2.5" EP MAIN *)	2
SSD SATA 6G 100GB Main 2.5" H-P EP	2
SSD SATA 6G 100GB MLC HOT P 2.5" EP MAIN *)	2
SSD SATA 3G 64GB SLC HOT P 2.5" EP MAIN *)	2
SSD SATA 3G 32GB SLC HOT P 2.5" EP MAIN *)	2
DOM SATA 3G 64GB Main N H-P	3

\*) EOL (製造終了)

データ冗長性やホットスペアおよびホットプラグ機能を持つ RAID ソリューションが長く利用されていますが、これはデータメディアの故障に対応して高可用性を保証するための技術的な方策といえます。

回復不能な読み取りエラーの統計的頻度を判断する指標がもう 1 つあります。Bit Error Rate (BER : ビットエラーレート) です。このエラー率は、一定のストレージ容量に対して予想されるビットエラー数を示しています。

次の表は、PRIMERGY サーバ向けの最新のデータメディアのビットエラーレートをまとめたものです。

ドライブクラス	ストレージタイプ	ビットエラーレート
エコノミック (ECO)	HDD	1 LBA / 10 <sup>15</sup> ビット <sup>1)</sup>
ビジネスクリティカル (BC)	HDD	1 LBA / 10 <sup>16</sup> ビット
エンタープライズ (EP)	HDD	1 LBA / 10 <sup>16</sup> ビット
	SSD	1 LBA / 10 <sup>16</sup> ビット
	PCIe-SSD	1 LBA / 10 <sup>20</sup> ビット
	DOM	n/a

<sup>1)</sup> LBA = Logical Block Address (論理ブロックアドレス)  
 10<sup>15</sup> ビット = 125 TB、10<sup>16</sup> ビット = 1.25 PB、10<sup>20</sup> ビット = 12.5 EB

この表を見ると、回復不能な読み取りエラーの確率に関する SSD の品質は、ビジネスクリティカル HDD やエンタープライズ HDD と同等です。このため、RAID システムの構成に使用する RAID レベルを選択する際には、同じセキュリティ基準に基づいて判断する必要があります。

このセキュリティ基準に関する PCIe-SSD の品質は、他より数桁上回っていますが、これには、全体的なフェールセーフの品質は反映されていません。

### 書き込み耐久性は何を意味しているのですか。

フラッシュメモリセルは消耗部品なので、SSD が実行できる書き込みジョブ数には限界があります。そのため、HDD のような全体の平均寿命だけでなく、書き込みジョブに関連した寿命指標が使用されています。これが書き込み耐久性です。

SSD メーカーは通常、SSD の書き込み耐久性を **PBW** (**P**etabytes **W**ritten) で示しています。以前は **TBW** (**T**erabytes **W**ritten) も使われていましたが、現在この指標が使用されているのは低容量の SSD だけです。**PBW** は、SSD に書き込める最大ペタバイト数を表し、

$$PBW = \frac{\text{容量 [GB]} \times (1 + \text{オーバープロビジョニング領域 [容量の \%]}) \times \text{フラッシュの耐久性 (プログラム/消去サイクル数)}}{\text{(ライトアンプリフィケーション要因} \times 10^6)}$$

この値が大きいほど高品質です。この値には、ランダムアクセスによる標準的な書き込み負荷として、SSD の典型的なライトアンプリフィケーションが考慮されています。つまり、この値はユーザーにとって実際に書き込めるペタバイト数を表しているのです。シーケンシャルアクセスパターンではライトアンプリフィケーションが低くなるため、PBW に示される値より書き込み耐久性が良くなる点にも留意してください。反対に、大きなライトアンプリフィケーションを許容するアクセス パターンは、その結果として書き込み耐久性が低下するともいえます。PBW はベンチマークとして捉えるようにしてください。

**DWPD** (**D**rive **W**rites **P**er **D**ay) は PBW から算出される指標です。一部の SSD メーカーは、SSD の書き込み耐久性を PBW ではなく DWPD で示しています。この指標の算出では、SSD の容量と既定の書き込み寿命も考慮されます。

$$DWPD = \frac{PBW \times 10^6}{\text{容量 (GB)} \times \text{既定の寿命 (年数} \times 365)}$$

この指標は、SSD がその既定寿命期間にわたって耐えられる 1 日あたりの平均書き込み負荷を示しています。通常、寿命には 5 年間つまり 1825 日が使用されます。DWPD 値が大きいほど高品質です。容量 400 GB、7.5 PBW の SSD の例を示します。

$$\frac{7.5 \times 10^6}{400 \times 5 \times 365} \approx 10 \text{ DWPD}$$

この SSD は、寿命期間中連続して、毎日その容量の 10 倍に相当する書き込み負荷への耐性があります。この計算では、SSD の容量が測定単位として使用されていることに注意してください。10 DWPD の 2 台の SSD のうち、容量の大きい方に書き込み負荷の高い処理を管理させることもできます。

容量 400 GB、10 DWPD の SSD :  $10 \text{ DWPD} \times 400 \text{ GB} \approx 3.6 \text{ TiB/日} \approx 44 \text{ MiB/秒}$

容量 200 GB、10 DWPD の SSD :  $10 \text{ DWPD} \times 200 \text{ GB} \approx 1.8 \text{ TiB/日} \approx 22 \text{ MiB/秒}$

1 台目の SSD は、容量が 2 倍であるだけでなく、許容書き込み負荷も 2 倍になるので、同じ書き込み負荷を与えた場合の耐久期間は 2 倍になります。

既定寿命の長さに注意することも重要です。メーカーによっては、一部の SSD モデルに 5 年より短い寿命を適用している場合もあります。寿命を短く設定すれば、DWPD 値が大きくなり、SSD が許容する 1 日あたりの書き込み負荷は大きくなりますが、これは短い期間での値です。容量 400 GB、7.5 PBW の SSD について、寿命を 5 年間ではなく 3 年間と想定した場合、計算結果は次のようになります。

$$\frac{7.5 \times 10^6}{400 \times 3 \times 365} \approx 17 \text{ DWPD}$$

$$17 \text{ DWPD} \times 400 \text{ GB} \approx 6.2 \text{ TiB/日} \approx 75 \text{ MiB/秒}$$

異なる SSD を比較するときに、最も重要性が高いのは PBW 指標です。異なる SSD の書き込み耐久性を DWPD 値に基づいて適切に比較するには、それぞれの容量が考慮されるとともに、寿命が同一である必要があります。これらを考慮しない場合、同じ書き込み負荷を与えても、10 DWPD の SSD の方が 15 DWPD の SSD よりも耐久期間がどうしても長くなる可能性もあります。

### DWPD 値と、1 日あたりの実際の書き込み負荷との関係を教えてください。

DWPD 値は、一定期間にわたって SSD が許容できる 1 日あたりの平均書き込み負荷の上限を表しています。1 日の実際の書き込み負荷がこれより小さければ、その分だけ、その SSD の書き込み耐久性は大きくなります。

実際の書き込み負荷が 500 GiB/日の場合について考えてみましょう。

「SSD SAS 6G 400GB MLC HOT PL 2.5 インチ EP PERF」を 10 DWPD で使用した場合、この SSD の書き込み耐久性は次のようになります。

$$\frac{\text{DWPD} \times \text{容量 (GB)} \times \text{既定の寿命}}{\text{書き込み負荷 (GiB/日)}}$$

$$\frac{10 \times 400 \text{ GB} \times 5 \text{ 年間}}{500 \text{ GiB}} \approx 37 \text{ 年}$$

この書き込み耐久性は、実際の使用限界を表しているわけではありません（一般的にほとんどのケースで同様です）。

### 現在使用している SSD を、書き込み耐久性が同等以上の SSD と交換したいと考えています。どうすれば、適切な SSD を選択できますか。

PBW 値が現在の SSD と同等以上の SSD であれば、DWPD 値が小さくても、書き込み耐久性は同等以上になります。[SSD の表](#)を参照してください。

## RAID 構成の選択は、SSD の書き込み耐久性に影響しますか。

原則的には影響します。

まず次の表のように、個々のデータメディアに対する書き込み負荷は SSD が組み込まれている RAID アレイのタイプによって異なります。

RAID レベル	SSD あたりの書き込み負荷
単一の SSD / JBOD	SSD および JBOD の書き込み負荷
RAID 0	RAID アレイの書き込み負荷 / ドライブ数
RAID 1	RAID アレイの書き込み負荷
RAID 1E	RAID アレイの書き込み負荷 × 2 / ドライブ数
RAID 5	RAID アレイの書き込み負荷 × 2 / ドライブ数
RAID 6	RAID アレイの書き込み負荷 × 3 / ドライブ数
RAID 10	RAID アレイの書き込み負荷 × 2 / ドライブ数
RAID 50	RAID アレイの書き込み負荷 × 2 / ドライブ数
RAID 60	RAID アレイの書き込み負荷 × 3 / ドライブ数

SSD 数の多い RAID アレイ設計では、SSD あたりの書き込み負荷が軽減されるので、書き込み耐久性の低い SSD で十分（状況によっては、そうしたタイプの SSD 方が経済的）という可能性もあります。

例：

「HD SAS 6G 146GB 15K HOT PL 2.5 インチ EP」が 7 台搭載された RAID 5 を SSD 使用の RAID 5 に交換します。論理ドライブの平均書き込み負荷は、今後 5 年間 6 TiB/日を超えないと考えられます。総容量は現在の構成と同等以上にする必要があります。使用できるドライブベイは 5 つまでです。

ソリューション 1：容量 400 GB の 3 台の SSD から成る RAID 5

SSD あたりの書き込み負荷：6 TiB/日 × 2 / 3 = 4 TiB/日 11 DWPD 以上必要  
→「SSD SAS 6G 400GB SLC HOT P 2.5 インチ EP PERF」（47 DWPD）

ソリューション 2：容量 400 GB の 4 台の SSD から成る RAID 5

SSD あたりの書き込み負荷：6 TiB/日 × 2 / 4 = 3 TiB/日 9 DWPD 以上必要  
→「SSD SAS 6G 400GB MLC HOT P 2.5 インチ PL EP」（10 DWPD）

高耐久性 SSD ではなく、主流の SSD を使用できるので、ソリューション 2 の方が明らかに経済的です。3 つ目のソリューションとして 10 DWPD で容量 440 GB 以上の SSD を 3 台使用するという方法もあります。

## 書き込み耐久性の期間が終了したらどうすればよいですか。

PRIMERGY サーバ向けに販売されている SSD と PCIe-SSD は、PBW の 95 %に達すると、書き込み操作の速度が大きく低下します。そのため、実際に PBW 値に到達することはほとんどなく、基本的に SSD は動作を維持します。システム管理者にとって、書き込みジョブの応答時間の増大が SSD 交換の目安となります。これは、各オペレーティングシステム固有のパフォーマンスカウンターを監視すれば確認できます。Windows オペレーティングシステムでは「書き込みバイト数/秒」のロギングによって、また Linux オペレーティングシステムでは iostat コマンドで、物理データメディアを監視できます。

PCIe-SSD の書き込み耐久性の監視とレポートには、管理ソフトウェアを使用できます。

## SSD へのアクセスがない場合、データが保持される期間はどの程度ですか。

SSD への給電が確保されていれば、内蔵 SSD の「スタティックウェアレベリング」機能によって、論理的には無限にデータが維持されますが、実際には SSD に障害が発生しない限りデータは維持されます（[MTBF と AFR](#) を参照）。

## 接続されていない SSD はどの程度の期間データを保持できますか。

SSD には、電源がオフになるとデータ保持機能が制約されるという特徴があります。たとえば、SSD をサーバから取り外し、バックアップとしてキャビネットに収納しておいた場合、保存情報を使用できるのは、長くても 10 年程度です。フラッシュテクノロジー（SLC/MLC）、それまでの使用状態（PBW）、周囲温度などの要因によって、保存期間がもっと短くなることもあります。SLC フラッシュストレージ型 SSD の最小保存期間は 6 ヶ月、MLC フラッシュストレージ型 SSD は 3 ヶ月です。

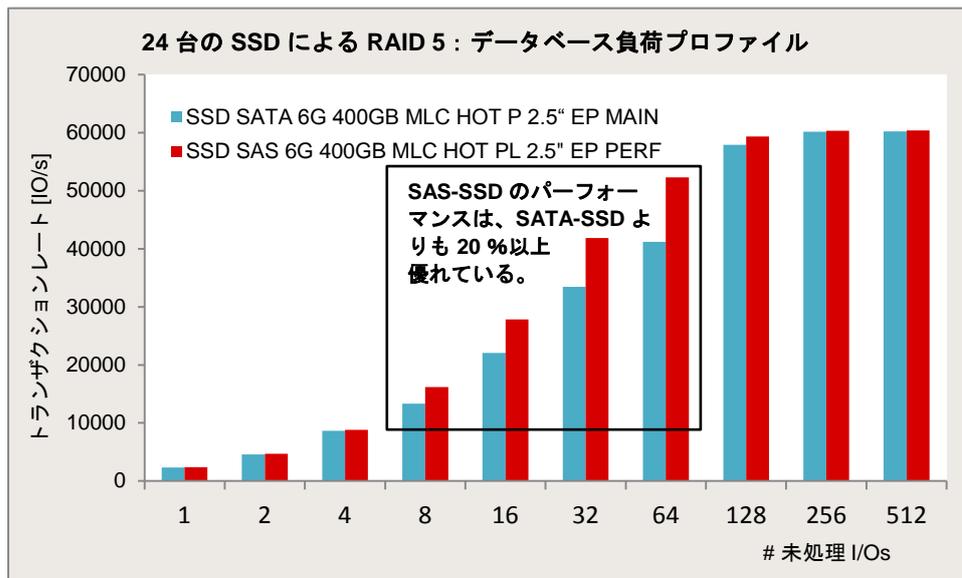
## インターフェースに関する FAQ

### SATA-SSD ではなく SAS-SSD を使用した方がよいのはどのような場合ですか。

SSD モデル間の違いが外部インターフェース (SATA / SAS) だけという場合もあります。スループットの上限だけを見れば、この両インターフェースに差異はありません。PRIMERGY サーバに搭載できる SSD の最大数も同じです。

しかし、SAS インターフェースには、SATA インターフェースより優れている点がいくつかあります。

- デュアルポート機能による冗長アーキテクチャ :  
SAS-SSD は 2 台の RAID コントローラーに接続できます。  
例 : PRIMERGY CX420 Cluster-in-a-box ソリューション
- エラー修正アルゴリズムが大幅に高性能
- パフォーマンス :  
一般的に、負荷が非常に低い場合や高い場合、SAS-SSD と SATA-SSD のパフォーマンスにほとんど差異はみられません。しかし、中程度の負荷の場合、負荷プロファイルによっては、SAS インターフェースの SSD のスループットが、SATA インターフェースの SSD より 20 % 以上高くなることもあります。ここでは、24 台の SSD で構成された RAID 5 に対するデータベース負荷プロファイルの例を示します。



データベース負荷プロファイル (トランザクション処理) : ランダム、読み取り 67%、書き込み 33%、ブロックサイズ 8 kB

## RAID に関する FAQ

### SSD を使用するには RAID が必要ですか。

SSD を使用する場合に RAID 構成を省略できるかどうかは、HDD 使用時になぜ RAID が必要なのかによって異なります。

- セキュリティ :  
SSD におけるフェールセーフの重要性は、HDD と同様です。HDD を使用した場合にデータ冗長性機能のある RAID 構成が必要であれば、SSD を使用しても RAID が必要となります。
- 容量 :  
最小容量値の大きい論理ドライブを実装する場合、通常、単一のデータメディアでは十分に対応できません。このような場合、RAID 構成は不可欠です。特に SSD の容量範囲は HDD より小さいので、RAID 構成の必要性は高くなります。
- パフォーマンス :  
HDD 使用時に RAID を使用する理由がパフォーマンスの問題だけなら、条件によっては RAID なしで SSD を使用できる場合もあります。ただし、これが可能なのは JBOD または RAID 0 として構成されている HDD を交換する場合だけです。

### HDD から SSD に変換する場合、RAID 構成について注意すべきことはありますか。

サーバのディスクサブシステムに HDD が使用されている場合、SSD ベースのディスクサブシステムへの変換時にまず必要なのは、既存の構成と交換できるような SSD 構成を定義することです。要求される書き込み負荷に対応するには一定数のストレージメディアが必要になります。これは容量とパフォーマンスによって決まります。SSD による構成は HDD による構成とは明らかに異なって見えます。これは、ストレージメディア数だけでなく、使用する RAID レベルや、HDD RAID アレイのキャッシュとして SSD を使用する場合にも関係する可能性があります。この点については、『[RAID Controller Performance](#)』も参照してください。さらに、HDD の場合とは異なり、RAID アレイ内の SSD の数を増やすと、SSD の寿命は大幅に伸びます。

### SSD または PCIe-SSD を使用する場合の推奨 RAID レベルを教えてください。

使用する RAID レベルを選択する際には、セキュリティの問題を考慮する必要があります。SSD に適用するセキュリティ基準は、基本的には HDD と同じです。MTBF の回数と故障率も、SSD と HDD の間で基本的に差異はありません。データメディア故障後の RAID アレイの再構築時間も同様です。RAID アレイ内の SSD のパフォーマンスが HDD を最も上回るのは、ブロックサイズの小さいランダム負荷プロファイルの場合です。ブロックサイズが大きくなると、この優位性は低下します。また、シーケンシャルアクセスでは差異はあるとしてもごくわずかになります。その場合、SSD とエンタープライズ HDD のどちらが優れているかは、RAID レベル、使用するデータメディア数、そしてコントローラーによって違ってきます。したがって、可能であれば、HDD にも SSD にも同じ RAID レベルを使用することを推奨します。ただし、現在の SSD は、一部の HDD タイプほど大容量ではないので、HDD RAID アレイを SSD に交換する場合、データメディア数を多くせざるを得ない場合もあります。例えば、現在、HDD による RAID 1 を SSD に代えるには、RAID 5 または RAID 10 を使用する必要があります。

### HW-RAID と SW-RAID

SSD は高性能であるため、ソフトウェア RAID を使用する場合は、パフォーマンスが高くなるほどプロセッサの負荷も増大します。この点については、HDD 使用時よりも十分に考慮する必要があります。SSD を使用すると、HDD 使用時よりもサーバプロセッサの負荷がかなり大きくなりますが、これは RAID コントローラーを使用することで緩和できます。ただし、RAID コントローラーは [TRIM](#) をサポートしていないので注意が必要です。

## SSD を使用する場合の推奨コントローラーを教えてください。

推奨コントローラーは、負荷プロファイルと、スループット、I/O レート、アクセス時間の要件によって異なります。これらは、容量とデータ冗長性の要件とともに、SSD の必要最小数やコントローラーの選択にも大きく影響します。これについての詳細は、ホワイトペーパー『[RAID Controller Performance](#)』を参照してください。

## キャッシュ付きの RAID コントローラーが必要なのはどのような場合ですか。

一定のライト比率で負荷範囲が低いまたは中程度のランダムアクセスでは、どのような RAID 構成でも、コントローラーキャッシュはプラスの影響を及ぼします。4 台以上の SSD による RAID 構成でシーケンシャルリードおよびライトアクセスを実行する場合も同様です。

## SSD を使用する場合、ホットスペアは必要ですか。

従来型 SSD の場合、原則的には必要といえます。もちろん、SSD には可動部分はありませんが、それでも故障することはあります。故障率 ([MTBF](#) を参照) はエンタープライズ HDD と同等です。

## さまざまな用途における SSD と HDD に関する FAQ

### HDD と SSD のパフォーマンスの違いはどの程度ですか。

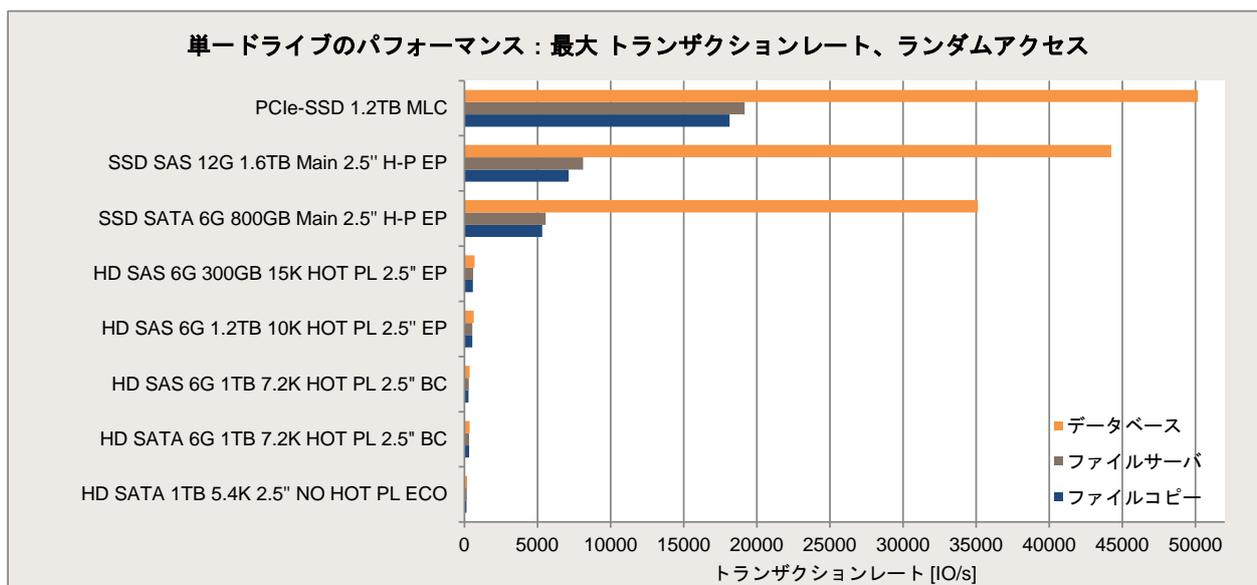
HDD では、リード/ライトヘッドの位置決めにかかる時間が大きな要因となります。ランダムアクセス、特に小さいブロックサイズを使用する場合には時間がかかります。そのため、HDD では、データベースの典型的なアクセス パターンにおけるスループット (MB/s) は、最大でも純粋なストリーミングのスループットの 3 % です。

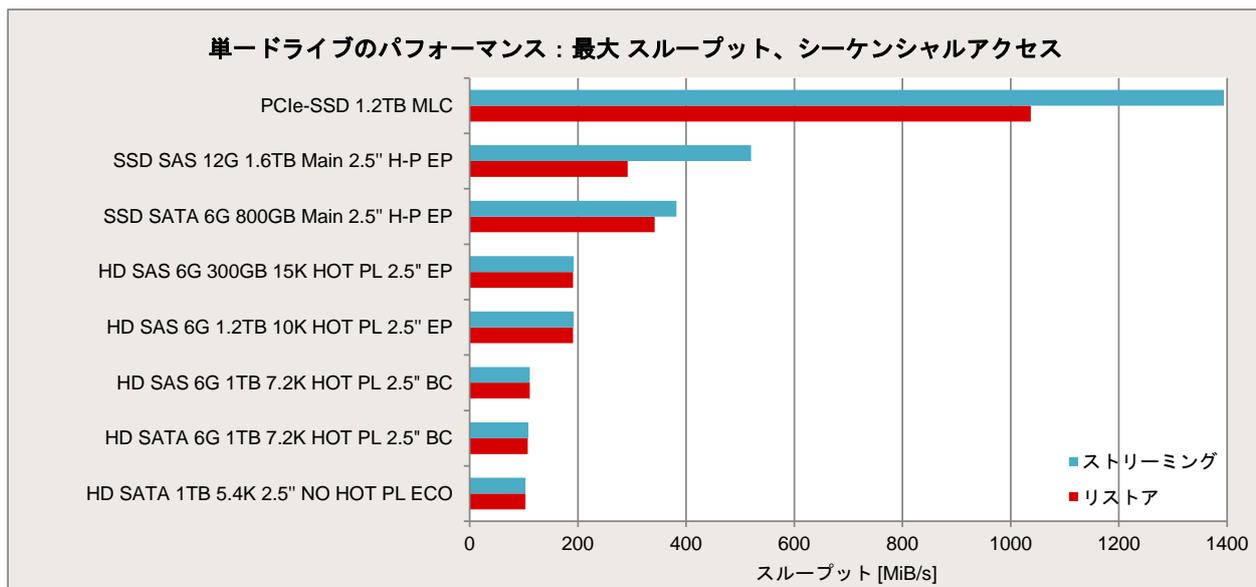
SSD を使用する場合は、この問題は技術的な理由で根本的に違ってきます。同じブロックサイズのランダムアクセス時のスループットは、シーケンシャルアクセスと比べてそれほど低下しません。典型的なデータベースアクセスパターンでも、SSD はストリーミングスループットの約 3/4 を達成します。

つまり、SSD のパフォーマンスの優位性は、特にランダムアクセスパターンの場合に明らかです。ブロックサイズが小さければ違いはさらに顕著になります。これに比べれば、エコノミック HDD とエンタープライズ HDD の間のパフォーマンスの差異は、ほとんど無視できるほど小さいものです。

以下の 2 つのグラフは、すべてのドライブカテゴリからデータメディアを選択し、それぞれの差異を示したものです。次の表に示されている標準の負荷プロファイルを使用しました。

標準負荷プロファイル	アクセス	アクセスの種類		ブロックサイズ [KiB]	アプリケーション
		リード	ライト		
ファイルコピー	ランダム	50 %	50 %	64	ファイルのコピー
ファイルサーバ	ランダム	67 %	33 %	64	ファイルサーバ
データベース	ランダム	67 %	33 %	8	データベース (データ転送) メールサーバ
ストリーミング	シーケンシャル	100 %	0 %	64	データベース (ログファイル)、 データバックアップ、 ビデオストリーミング (一部)
リストア	シーケンシャル	0 %	100 %	64	ファイルのリストア





パフォーマンスの詳細については、『[RAID Controller Performance](#)』、『[Performance Report PCIe-SSDs ioDrive®2](#)』のほか、『[PRIMERGY Performance Reports](#)』のディスク入出力のセクションを参照してください。

その他ご不明の点については、PRIMERGY Performance and Benchmarks (<mailto:primergy.benchmark@ts.fujitsu.com>) までお問い合わせください。

## このホワイトペーパーでは私の疑問に対する答えが見つかりません。なぜですか。

著者はすでにあなたの質問を把握し、それを適切に反映するために改訂しようとしているかもしれません。

問い合わせ先 :

PRIMERGY Product Marketing (<mailto:Primergy-PM@ts.fujitsu.com>) または

PRIMERGY Performance and Benchmarks (<mailto:primergy.benchmark@ts.fujitsu.com>)

## 関連資料

### PRIMERGY システム

<http://jp.fujitsu.com/primergy>

### PRIMERGY コンポーネント

このホワイトペーパー :

 <http://docs.ts.fujitsu.com/dl.aspx?id=78858d6c-4c0f-479a-8ceb-705fe1938f4e>

 <http://docs.ts.fujitsu.com/dl.aspx?id=222ef5d3-0d3b-428a-9472-e8ed1ca1c8b9>

 <http://docs.ts.fujitsu.com/dl.aspx?id=1d8b7d65-e5f4-4a99-8e7b-f47c74ccc85e>

ホワイトペーパー :

Hard disk drives or solid state disk drives for servers – what is more suitable? (英語)

<http://docs.ts.fujitsu.com/dl.aspx?id=94cf1265-15d9-4d91-bbba-0345f37eb74b>

SSD 書き込み耐久性のソリューション別評価

<http://docs.ts.fujitsu.com/dl.aspx?id=8325f1fc-f98c-4467-a3aa-98f6c633143f>

ディスク I/O パフォーマンスの基本

<http://docs.ts.fujitsu.com/dl.aspx?id=35801735-a223-491a-a879-43f506444366>

パフォーマンスレポート PCIe-SSDs ioDrive®2

<http://docs.ts.fujitsu.com/dl.aspx?id=91b93b35-7511-45e4-a5be-ccde9219772f>

RAID コントローラーのパフォーマンス

<http://docs.ts.fujitsu.com/dl.aspx?id=e34159fa-0196-4a01-99ff-8792b5f644eb>

### PRIMERGY のパフォーマンス

<http://jp.fujitsu.com/platform/server/primergy/performance/>

## お問い合わせ先

### 富士通

Web サイト : <http://jp.fujitsu.com/>

### PRIMERGY Product Marketing

<mailto:Primergy-PM@ts.fujitsu.com>

### PRIMERGY のパフォーマンスとベンチマーク

<mailto:primergy.benchmark@ts.fujitsu.com>

© Copyright 2013-2014 Fujitsu Technology Solutions。Fujitsu と Fujitsu ロゴは、富士通株式会社の日本およびその他の国における登録商標または商標です。その他の会社名、製品名、サービス名は、それぞれ各社の登録商標または商標です。知的所有権を含むすべての権利は弊社に帰属します。製品データは変更される場合があります。納品までの時間は在庫状況によって異なります。データおよび図の完全性、事実性、または正確性について、弊社は一切の責任を負いません。本書に記載されているハードウェアおよびソフトウェアの名称は、それぞれのメーカーの商標等である場合があります。第三者が各自の目的でこれらを使用した場合、当該所有者の権利を侵害することがあります。

詳細については、<http://www.fujitsu.com/fts/resources/navigation/terms-of-use.html> を参照してください。

2014-03-31 WW JA