

# HP ProLiantサーバー用ソリッド ステート ストレージ テクノロジー

## 第2版

### 技術概要

はじめに .....	2
フラッシュメモリテクノロジー .....	2
シングルレベルセルおよびマルチレベルセルNANDフラッシュ .....	2
NANDアーキテクチャー .....	3
フラッシュメモリを使用したソリッドステートドライブの設計 .....	4
SSDの耐久性を改善するウェアレベリング .....	4
NANDのオーバープロビジョニング .....	4
HP ProLiantサーバー用のHPソリッドステートドライブ .....	5
HPサーバーSSDのパフォーマンス .....	5
HPサーバーSSDの信頼性 .....	5
IOアクセラレータ・新しいブロックストレージデバイス .....	6
HP PCIe IOアクセラレータのモデル .....	7
HP PCIe IOアクセラレータのパフォーマンス .....	7
まとめ .....	8
詳細情報 .....	9
ご意見の送り先 .....	9



## はじめに

HPは回転式磁気メディアに基づく従来のディスクドライブに加えて、フラッシュメモリに基づくソリッドステートストレージデバイスを提供しています。ソリッドステートドライブ(SSD)はディスクドライブと互換性があり、従来のSATA/SASディスクコントローラーと組み合わせて使用できるので、このような新しいデバイスの中で最も普及しています。また、HPはHP ProLiantサーバー用に特定アプリケーション向けにパフォーマンスを大幅に改善する、フラッシュに基づく新しい種類のストレージデバイスも導入しています。この技術概要では、ソリッドステートストレージについて、また、この新しいテクノロジーを使用して開発されている高性能な新製品について説明します。

## フラッシュメモリテクノロジー

ソリッドステートドライブの多くは、電氣的に消去および再プログラムできる不揮発性コンピューターメモリであるフラッシュメモリテクノロジーを使用しています。フラッシュには、NORフラッシュとNANDフラッシュの2つの種類があります。NORフラッシュとNANDフラッシュはともに、「セル」と呼ばれるフローティングゲートトランジスターのアレイに情報を保存します。しかし、この2つは、セルアレイの接続方法とアクセス方法が異なります。NORフラッシュメモリのセルはビット線に並列に接続され、セルを個別に読み取りおよびプログラムできます。NANDフラッシュメモリのセルは直列に接続され、セルをグループとしてのみ読み取りおよびプログラムできます。

NANDアーキテクチャーでは、同等のNORメモリの約2倍にあたる密度のメモリアレイをより低いコストで製造できます。そのため、大多数のデバイスではNANDフラッシュメモリが使用されています。

## シングルレベルセルおよびマルチレベルセルNANDフラッシュ

NANDフラッシュテクノロジーには主に2つの種類があります。

- シングルレベルセル(SLC)テクノロジーでは各セルに単一レベルの電荷を保存することで1ビットの情報を表現します。
- マルチレベルセル(MLC)テクノロジーでは各セルに4種類の電荷状態のいずれかを保存します。そのため、各セルが2ビットの情報を表現できるので実質的にストレージ密度が2倍になります。

マルチレベルセルテクノロジーを使用するNANDフラッシュメモリは、消費者向け製品における主要フラッシュテクノロジーとして急速に広まりました。しかし、SLCに比べ、MLCテクノロジーにはサーバーストレージ用途として要求される性能や信頼性の面で以下のような特性の違いがあります(表1)。

- セルの状態を区別するための差が小さいため内部エラー率が比較的高く、エラーを訂正するために必要なECCメモリが増える。
- 最大プログラム/消去サイクル数の面で寿命が短い。
- 読み取り/書き込み(プログラム)パフォーマンス、特に書き込みのパフォーマンスがSLCより低い。

表1. SLCフラッシュとMLCフラッシュの主な特性

	SLCフラッシュ	MLCフラッシュ
ランダムアクセス	25マイクロ秒	60マイクロ秒
シーケンシャルアクセス	50ナノ秒	30ナノ秒
ページのプログラム(書き込み)	200マイクロ秒	800マイクロ秒
最大プログラム/消去サイクル数	100,000回(1ビットECC)	5000~10,000回(4ビットECC)

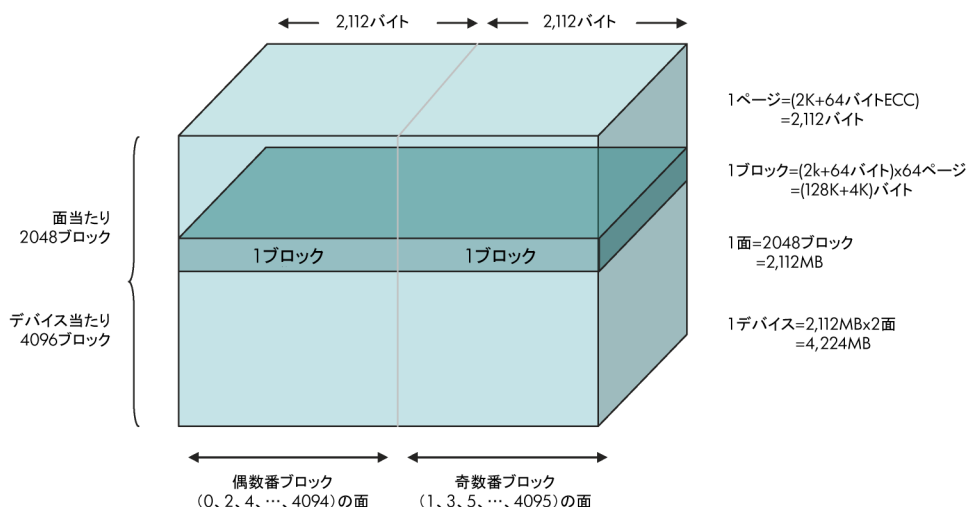
表1に示すように、MLC NANDフラッシュは、SLCフラッシュに比べて読み取りおよび書き込みパフォーマンスが劣ります。より重要なことは、SLCフラッシュのプログラム/消去ライフサイクル(耐久性とも呼ばれます)がMLCフラッシュの10~20倍であることです。SLC NANDの優れたパフォーマンスと信頼性は、ワークロードが制約されない環境でソリッドステートドライブを利用する場合に適しています。一方、書き込み頻度は低く読み取り比率の多いアプリケーション環境では、MLCベースのSSDも使用されます。

## NANDアーキテクチャー

NANDフラッシュメモリアレイはページとブロックで構成されます。ページは最小単位です。ページサイズはNAND実装によって異なりますが、一般に2KB、4KB、または8KBです。ページはブロックにまとめられ、一般に1ブロックは64ページで構成されます。ここでは、これらの単位をSATA/SASインターフェイスの512バイト論理ブロックと区別して、「NANDブロック」と呼びます。

また、SLC NANDはデバイスを奇数ブロックと偶数ブロックからなる2枚の物理的な面に分割する2面アーキテクチャーで実装されることがあります。2面フラッシュでは、2つのページを並列に読み取りおよびプログラムできるため、NANDのパフォーマンスが向上します。また、2つのブロックを並列に消去できます。図1にブロック当たり64の2Kページで構成される4GBのSLC NANDアーキテクチャーを示します。NANDアーキテクチャーは現在も急速に発展を続けており、8Kページが一般的となり4面設計が登場しつつあります。

図1. NANDメモリの構成



NANDフラッシュには情報の書き込みおよび取得のための特定のプロトコルがあります。読み取りまたは書き込みが可能な最小単位はページです。ディスクドライブと異なり、既存のデータを含んでいるページは新しいデータで直接更新することができません。最初に既存のデータが消去されます。NANDメモリでは、通常64ページまたは128ページで構成されるNANDブロック全体を消去することしかできません。NANDフラッシュを使用するストレージデバイスの重要な課題の1つが、消去と書き込みのサイズの差に効果的に対処することです。表2に、これらの基本的なNAND操作とその実行時間を示します。

表2. SLC NANDフラッシュの操作

操作	最小実行時間
ランダムなページ読み取り	25マイクロ秒
ページのプログラム(書き込み)	200マイクロ秒
ブロック消去	1500マイクロ秒

表2に示すようにNANDフラッシュの書き込みには読み取りよりも時間がかかります。ページプログラム操作には、ランダムなページ読み取りの8倍の時間がかかります。ブロック消去操作は実行頻度は低いですがページプログラム操作の7倍の時間がかかります。この実行時間の差に対しては、多くの高レベル戦略で対応策が講じられていますが、USBドライブやより高度なソリッドステートドライブなど、NANDに基づくすべてのストレージデバイスで読み取りパフォーマンスが書き込みパフォーマンスより優れている主な理由です。

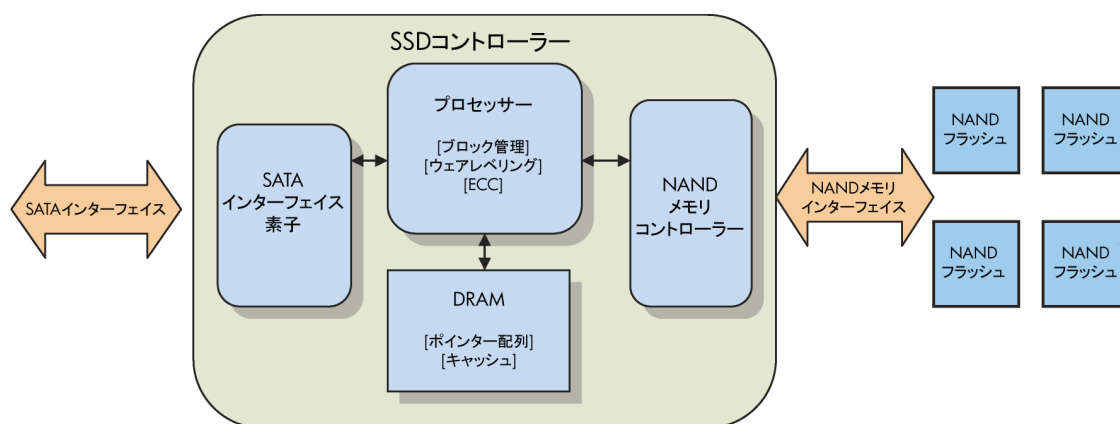
## フラッシュメモリを使用したソリッドステートドライブの設計

NANDに基づくソリッドステートドライブには、次のようなタスクを実行するドライブコントローラーサブシステムが必要です。

- エラー処理やブロック管理を含むNANDメモリの読み取り/書き込み操作の管理
- 管理アルゴリズムとRAMベースキャッシュを使用したNANDフラッシュのパフォーマンスの強化
- NANDメモリに対する書き込み/消去サイクルを最小限に抑えるアルゴリズムによるSSDの耐久性(寿命)の最大化
- NANDの読み取り/書き込みインターフェイスと、ホストに必要なインターコネク(通常SASまたはSATA)間の変換

図2は、典型的なSATA SSDの機能図です。NANDフラッシュメモリを管理し、標準的なSATAストレージインターフェイスをホストサーバーに提供するために必要なすべての操作ロジックを含むSSDコントローラーを示しています。

図2 典型的なSATAソリッドステートドライブの機能図



### SSDの耐久性を改善するウェアレベリング

ウェアレベリングはNANDに基づくSSDの耐久性を改善するために使用される設計手法の1つです。NANDに基づくSLCフラッシュでは、100,000回ほどの書き込み/消去サイクルを実行すると寿命となるため、SSDでは必要以上にNANDブロックの消去や書き換えを行わないようにする必要があります。しかし、アプリケーションによっては、SAS/SATAデバイスの一部の論理SCSIブロックの更新や書き換えを頻繁に行う必要がある場合があります。ウェアレベリングは、論理SCSIブロックをNANDアレイ内の異なる物理ページに継続的にマッピングし直すことで、この問題を解決します。ウェアレベリングによって消去と書き換えをメディア全体に均等に分散することで、SSDの耐久性を最大に高められます。SSDのパフォーマンスを最大化するために、SSDコントローラーは、論理対物理マッピングを高速なDRAM内のポインター配列として保持します。また、この情報はNANDフラッシュアレイ内のメタデータ領域にも自動的に保持されます。これらの手法により、SSDへの電源供給が予期せず途切れた場合でもSSDでマッピングを再構築することができます。

### NANDのオーバープロビジョニング

設計技術者はデバイス上のNAND容量をオーバープロビジョニングすることでSSDの耐久性とパフォーマンスを改善することができます。オーバープロビジョニングにより、より多くのNANDブロックに書き込みと消去を分散できるため、SSDの耐久性を改善できます。また、SSDコントローラーがページ書き込みやNANDブロック消去を管理するためのバッファ領域が増えるため、SSDのパフォーマンスも向上します。ハイエンドのSSDでは表示されているストレージ容量の25%ものNANDメモリをオーバープロビジョニングする場合があります。

## HP ProLiantサーバー用のHPソリッドステートドライブ

HPは2008年にサーバー用の最初のSSDを導入しましたが、ホットプラグ対応でなく特定のBladeServer環境のみを対象としたSSDでした。2009年にはホットプラグに対応した従来のドライブキャリアでのSSDを導入しました。この3Gbps SATA SSDは従来のミッドラインSATAディスクドライブが使用できるほとんどのHP ProLiantサーバーで使用できます。サーバー用SSDはPCベースのソリッドステートドライブと異なり、サーバーストレージデバイス向けの高い標準を満たしています。同時に、SSDの特性であるパフォーマンスと信頼性を備えています。

### HPサーバーSSDのパフォーマンス

ディスクアクセスタイム(レイテンシともいいます)は、ドライブからデータを取り出すために必要な合計時間であり、従来のディスクドライブではパフォーマンスを左右する要素です。ディスクドライブのレイテンシはシークタイム、回転待ち時間、転送時間の合計です。

SSDではシークタイムや回転待ち時間がありません。レイテンシはメモリアクセスタイムと転送時間にコントローラーのオーバーヘッドを合わせたものになります。以上の事実とNANDフラッシュの一般的な特徴から、次のようなことがいえます。

- SSDでは、NANDプログラム(書き込み)操作が相対的に遅いため、読み取り操作の方が書き込み操作より速い。
- SSDでは、各読み取り操作でシークタイムや回転待ち時間がないため、SSDのランダム読み取りはディスクドライブのランダム読み取りより速い。

表3に32GB SFF HPサーバーSSDと15K SASハードディスクドライブ(HDD)のパフォーマンスの比較を示します。シーケンシャル操作のパフォーマンスでは大差はありませんが、ランダム操作のパフォーマンスはSSDの方が大幅に優れています。SSDのランダム読み取りパフォーマンスはHDDのパフォーマンスの50倍を上回ります。

表3. 典型的なSSDとHDDのパフォーマンスの比較(実際のパフォーマンスの値とは異なります)

	HP 3Gbps SATA SSD	HP SFF 15K SAS HDD
ランダム読み取り(4KB)	18,000 IOPS	340 IOPS
ランダム書き込み(4KB)	3000 IOPS	285 IOPS
シーケンシャル読み取りスループット(64KB)	220MB/秒	105MB/秒
シーケンシャル書き込みスループット(64KB)	120MB/秒	150MB/秒

### HPサーバーSSDの信頼性

信頼性はサーバーで使用するストレージデバイスを選択する際の重要な基準です。しかし、すべてのSSDがハードディスクドライブより優れた信頼性を備えているとは限りません。信頼性の高いSSDはNANDフラッシュアレイを管理でき、また、以下のようなNANDエラーによって引き起こされる問題を是正できるようなコントローラー設計を備えている必要があります。

- 読み取り失敗
- プログラム(書き込み)失敗
- ホットチャージインジェクション(不良セルのビットが反転する)

サーバー用のHPソリッドステートドライブはサーバー環境に必要な高い信頼性を実現する各種メカニズムを採用しています。以下にそのメカニズムを示します。

- 耐久性に優れたSLC NANDテクノロジー
- 寿命を延ばすためのNANDメモリのオーバープロビジョニング
- ウェアレベリングとブロック管理
- NANDエラーの頻度を大幅に減らす、読み取りと書き込みのアルゴリズム
- エンドツーエンドのデータバスエラー検出
- 不意の停電からの保護

SLCを採用したサーバー用のHPソリッドステートドライブは、これらのテクノロジーを通じて現在のサーバー用HPエンタープライズディスクドライブと同程度の信頼性を実現します。

HPサーバーSSDは高温の環境、衝撃と振動の多い環境など、従来のディスクドライブには適していない条件でも上記の機能を発揮できるため、特定のアプリケーションにとってはきわめて信頼性の高いデバイスといえます。表4にHPサーバーSSDとHP SFF SASエンタープライズドライブの動作範囲の比較を示します。

**表4. SSDとHDの動作範囲の比較**

	HP 3Gbps SATA SSD (2009年発売)	HP SFF 15K SAS HDD
動作温度	0~55°C	10~35°C
動作時の衝撃	1500g(0.5ミリ秒の半正弦波)	30g(2ミリ秒の半正弦波)
振動	20g(ピーク) 10~2000Hz	1.5g(RMS) 10~500Hz
消費電力(動作時)*	2ワット未満	8~9ワット

\*注: 今後発売予定のデバイスではパフォーマンス向上のため、消費電力が増加する場合があります。

## IOアクセラレータ - 新しいブロックストレージデバイス

IOアクセラレータはフラッシュメモリテクノロジーに基づく新しいクラスのストレージ製品です。IOアクセラレータは最新のすべてのストレージと同様、標準的なブロックレベルのストレージデバイスであり、他のストレージボリュームと同じようにアクセスできます。ただし、SSDと異なり接続互換性のあるストレージコンポーネントではなく、標準のドライブコントローラーやSASまたはSATAチャネルを通じてアクセスするものではありません。IOアクセラレータはストレージデバイスであると同時にコントローラーとしての機能も果たし、専用ドライバーが必要で、PCIeバス経由でブロック/Oを直接提供します。後述するように、特定のアプリケーション環境で使用する際に明確なパフォーマンス上の利点があります。IOアクセラレータのアーキテクチャー

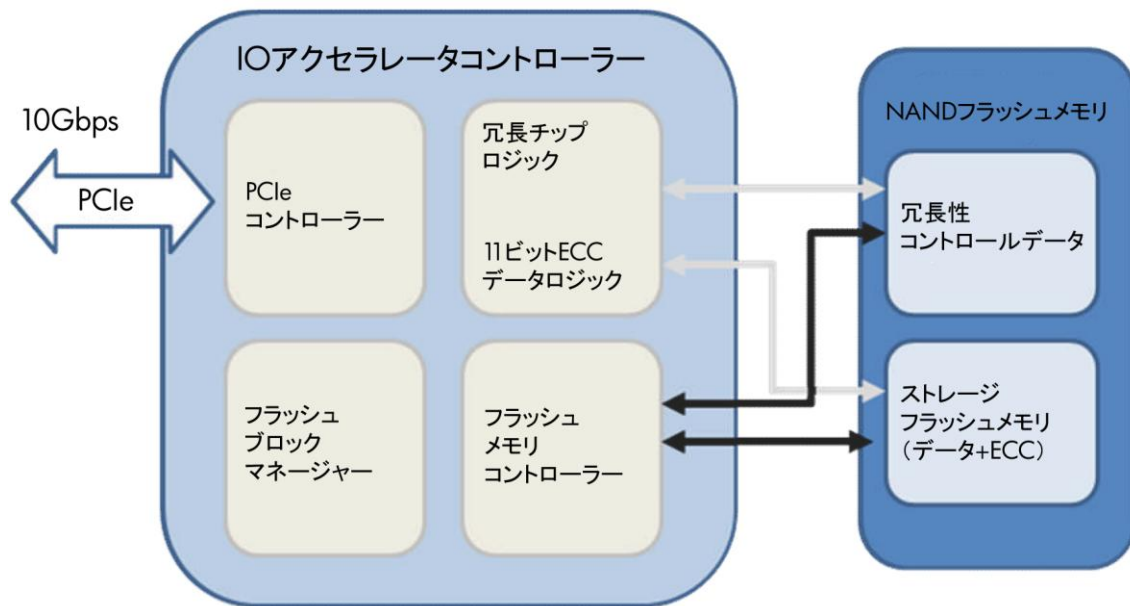
IOアクセラレータはSSDと同じ基本的なNANDメモリおよびNANDメモリコントローラーテクノロジーを使用し、フラッシュメモリとの間でデータの格納と取得をローレベルで実行しますが、以下の点ではSSDと異なります。

SSDは図2に示したようにオンボードプロセッサを使用して、NAND読み取り/書き込みインターフェイスと、Smartアレイホストコントローラーに提示されるストレージブロックインターフェイスとの間の変換を実行します。ブロックデータは3Gbpsまたは6Gbpsリンク経由でコントローラーに移動してから、PCIeバス経由で呼び出し側のアプリケーションに提供されます。

これに対してIOアクセラレータは、コントローラーとして機能するストレージデバイスであり、デバイスドライバーが必要なPCIeカードです。IOアクセラレータでは標準のブロックレベル/OをNAND読み取りおよび書き込みに変換するすべてのロジックが、アクセラレータのデバイスドライバーに含まれています。ウェアレベリング、エラー訂正、不良ブロック管理などのNAND管理機能についても同様です。このアーキテクチャーによりIOアクセラレータでは、サーバーCPUの大幅に広い帯域幅とマルチコアプロセッシング能力を利用して、ブロックストレージスループットの大幅な向上と、オンボードプロセッサでは実現できない低いレイテンシを実現できます。



図3. HP IOアクセラレータのアーキテクチャー



また、IOアクセラレータではSSDより広く平坦なNANDセルアレイを使用します。典型的なSSDではNANDチャンネルが4つであるのに対し、HP PCIe IOアクセラレータはNANDアレイで8～12のNANDチャンネルを使用します。このアーキテクチャーにより、HPのIOアクセラレータはより多くのNANDアクセスを並列実行できるので、典型的なSSDより大きなパフォーマンスの改善を実現します。

最後に、HP PCIe IOアクセラレータのアーキテクチャーにより、ソリッドステートメモリとPCIeバスとの間でボトルネックになる場合がある中間段階がなくなるため、最大限にパフォーマンスが活かされます。SASチャンネルとSmartレイアウトコントローラーフロントエンドがないため、スループットを物理的に制限するのはPCIeバスの帯域幅のみです。

## HP PCIe IOアクセラレータのモデル

PCIe IOアクセラレータにはioDriveとioDrive Duoの2種類があります。ioDriveは160GBのSLC NANDメモリまたは320GBのMLC NANDを搭載するロープロファイルPCIe x4(4レーン)デバイスです。

ioDrive Duoはハードウェアレベルでは、2つのNANDストレージデバイスから構成されます。オペレーティングシステムやアプリケーションによってはソフトウェアRAID0を利用することで単一のストレージボリュームとして扱うこともできます。ioDrive Duoは、このアーキテクチャーと広帯域幅のPCIe x8(8レーン)での接続によってioDriveより優れたI/Oパフォーマンスを提供します。320GB ioDrive DuoはSLC NANDとioDrive Duoアーキテクチャーの広帯域幅を使用するため、すべてのモデルの中で最高のパフォーマンスを提供します。

## HP PCIe IOアクセラレータのパフォーマンス

HP PCIe IOアクセラレータ製品は、高いIOPSとスループットパフォーマンスおよび低いレイテンシが利点となるアプリケーション環境に対し、きわめて高速のブロックストレージパフォーマンスを提供します。以下に、このようなアプリケーション環境の例を示します。

- データベースおよびデータベース高速化
- Webサーバー
- ビデオ、レンダリング、アニメーション

表5にこれらのデバイスの主なパフォーマンス特性を示します。

表5. ioDriveのパフォーマンス

	HP 160 GB ioDrive	HP 320 GB ioDrive Duo	HP 640 GB ioDrive Duo
NANDの種類	SLC	SLC	MLC
シーケンシャル書き込みスループット(32kブロック)	670MB/秒	1.4GB/秒	1.0GB/秒
シーケンシャル読み取りスループット(32kブロック)	750MB/秒	1.5GB/秒	1.4GB/秒
IOPS(4kランダム読み取り)	116,000	185,000	122,000
読み取りアクセスレイテンシ	26マイクロ秒	50マイクロ秒	80マイクロ秒

## まとめ

NANDフラッシュメモリを使用するソリッドステートストレージ製品は、急速に発展する新しいクラスのHP ProLiant サーバー用製品です。今日、ソリッドステートストレージ製品は、従来の回転式メディアでは対応できないランダム IOPS パフォーマンスを必要とするアプリケーション環境で使用されます。HPソリッドステートドライブは、大幅に向上したI/Oパフォーマンスを提供できるとともに、Smartアレイストレージ環境およびハードウェアベースRAIDやホットプラグ対応ドライブを含むすべての機能とシームレスに統合できます。HP IOアクセラレータは、PCIeバス経由で直接動作することで最大限のストレージI/Oスループットを実現するように設計された、新しいクラスのブロックストレージデバイスです。



## 詳細情報

本書で説明した内容について詳しくは、以下の資料を参照してください。

リソースの説明	Webアドレス
HPソリッドステートストレージのWebページ	<a href="http://www.hp.com/jp/solidstate">http://www.hp.com/jp/solidstate</a>
HP ProLiantドライブ(ソリッドステートドライブを含む)	<a href="http://h50146.www5.hp.com/products/servers/proliant/storage/diskstorage.html">http://h50146.www5.hp.com/products/servers/proliant/storage/diskstorage.html</a>
HP PCIe IOアクセラレータのWebページ	<a href="http://h50146.www5.hp.com/products/servers/proliant/storage/io_acce/">http://h50146.www5.hp.com/products/servers/proliant/storage/io_acce/</a>
TPC-Hベンチマークを使用した、SSD、ioDrive、SAS回転式ドライブの比較	<a href="http://h20195.www2.hp.com/v2/GetPDF.aspx/4AA0-0248ENW.pdf">http://h20195.www2.hp.com/v2/GetPDF.aspx/4AA0-0248ENW.pdf</a> (英語)

## ご意見の送り先

本書についてご意見がありましたら、[TechCom@HP.com](mailto:TechCom@HP.com) (英語)までお送りください。

© 2010 Hewlett-Packard Development Company, L.P. 本書の内容は、将来予告なしに変更されることがあります。HP製品およびサービスに対する保証については、当該製品およびサービスの保証規定書に記載されています。本書のいかなる内容も、新たな保証を追加するものではありません。本書の内容につきましては万全を期しておりますが、本書中の技術的あるいは校正上の誤り、脱落に対して、責任を負いかねますのでご了承ください。本書に記載されている仕様は、予告なしに変更されることがあります。

