

RAIDコントローラー (オンボード/オプション)

SAS/SATA/NVMe

Type-B

本ガイドについて

本ガイドは、RAID コントローラーに関し、ユーザズガイドでは説明しきれない詳細な情報を記載しております。

目次

第 1 章 基礎知識編	4
1.1 RAID	4
1.1.1 スプリットシークによる高速化	4
1.1.2 パックによる大容量化	4
1.1.3 冗長構造による高信頼性	4
1.2 RAID コントローラー	5
1.2.1 PCI カードタイプ	6
1.2.2 メザニンカードタイプ	7
1.2.3 オンボードタイプ	8
1.3 PCI 規格	10
1.3.1 PCI Express	10
1.3.2 PCI ボードのサイズ	10
1.4 物理デバイスのインタフェース	11
1.4.1 SAS 規格	11
1.4.2 SATA 規格	11
1.4.3 SAS/SATA 比較	11
1.4.4 SAS/SATA の奨励用途	12
1.4.5 セクタサイズ	12
第 2 章 機能編	13
2.1 RAID システム構築機能	14
2.1.1 ディスクアレイ (パック)	14
2.1.2 ディスクアレイ (パック) の構成ルール	14
2.1.3 RAID の種類 (RAID レベル)	15
2.2 初期化	22
2.2.1 初期化方法	22
2.3 再構築 (リビルド)	22
2.3.1 ホットスペア (スタンバイ) リビルド	22
2.3.2 ホットスワップリビルド	22
2.3.3 リビルド時間目安	22
2.3.4 スペアドライブ	23
2.4 表面スキャン	24
2.5 電力モード	24
2.6 キャッシュ機能 Flash-Backed write cache (FBWC)	25
2.6.1 無効	25
2.6.2 有効	25
2.6.3 バッテリー	25
2.6.4 フラッシュバックアップユニット	26
2.7 アレイ変換と論理ドライブの移行	26
2.7.1 アレイの拡張	26
2.7.2 アレイの移動	26
2.7.3 アレイの交換	26
2.7.4 アレイの縮小	26
2.7.5 アレイのミラー化	27
2.7.6 アレイの修復	27
2.8 論理ドライブの移行	28
2.8.1 論理ドライブの拡大	28
2.8.2 RAID レベルの変更	28
2.8.3 ストライプサイズの移行	28
2.8.4 移行の優先度	28

第3章 ハードウェア編.....	29
3.1 RAID コントローラー製品一覧.....	29
3.2 各 RAID コントローラーの仕様.....	30
3.3 バックアップユニットの仕様.....	30
3.4 フラッシュバックアップユニットの仕様.....	30
3.5 物理デバイス選定における確認事項.....	30
3.6 注意事項.....	30
第4章 ソフトウェア編.....	31
4.1 RAID コントローラーのソフトウェア.....	31
4.2 オフラインユーティリティ.....	31
4.3 オンラインユーティリティ.....	31
4.4 RAID コントローラーのソフトウェアでの注意事項.....	32
第5章 運用編.....	33
5.1 性能比較.....	33
5.1.1 性能比較.....	33
5.2 RAID レベルの比.....	33
5.3 RAID システムの構築.....	34
5.4 安定運用のために.....	35
5.4.1 表面スキャンの実施.....	35
5.4.2 RAID システム管理ユーティリティ+ESMPRO の利用によるアラート.....	36
5.4.3 RAID システム管理ユーティリティの使用について.....	36
5.4.4 RAID コントローラー用ドライバ、RAID システム管理ユーティリティのアップデート.....	36
5.4.5 RAID 構成物理デバイス台数の設定による保守運用性の向上.....	37

第1章 基礎知識編

1.1 RAID

サーバを構成する部品の中でハードディスクドライブは機械的な動作を伴う為に非常にデリケートです。その上 CPU やメモリなど他の構成部品と比べ桁違いに動作速度が遅いのも特徴です。RAID (Redundant Array of Independent Disks)とは複数台のハードディスクドライブを用いて、I/O 処理を分散する事で高速化し、データとそのパリティを分散して格納する事で大容量化・高信頼性を確保する技術です。

1.1.1 スプリットシークによる高速化

ハードディスクドライブは機械的な動作を伴うために、CPU やメモリに比べると桁違いに遅くなります。しかし、速度の遅い機械的な動作でも複数台のハードディスクドライブを用いて同時におこなう事(スプリットシーク)でファイルの I/O 性能を向上させることができます。

1.1.2 パックによる大容量化

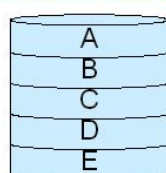
複数台のハードディスクドライブを単一ドライブとして認識させる(パックする)ことで、大容量のドライブを構成することができます。10TB のハードディスクドライブを一台準備することは現時点では難しいですが、2TB のハードディスクドライブを5台準備することで10TBのドライブを構成することができます。

1.1.3 冗長構造による高信頼性

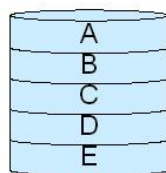
格納されているデータとそのパリティを保存することで、論理ドライブに冗長性を持たせることができます。冗長構造を持つ論理ドライブを構成することで、ハードディスクドライブに障害が発生してもシステムを停止せずに復旧作業をおこなうことができます。

(例) 1 台のHDDに障害が発生した場合

ミラーリング



HDD1

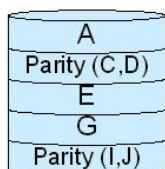


HDD2

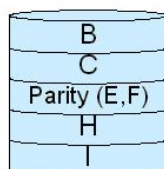
HDD1より全てのデータを読み出すことが可能

データ=A,B,C,D,E...

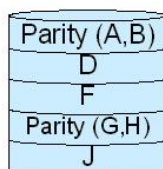
パリティ



HDD1



HDD2



HDD3

パリティより算出 D,F,J

HDD1より A,E,G

HDD2より B,C,H,I

データ=A,B,C,D,E,F,G,H,I,J...

1.2 RAID コントローラー

RAID コントローラーは複数のハードディスクドライブにて構成される RAID システムに対し、パリティ計算やデータ読み出し/書き込み等の処理を行う専用ハードウェアです。RAID コントローラーの機能を持ち PCI バスへ接続するオプションカードタイプと、RAID 機能を提供するチップをマザーボード上に直接実装するオンボードタイプがあります。

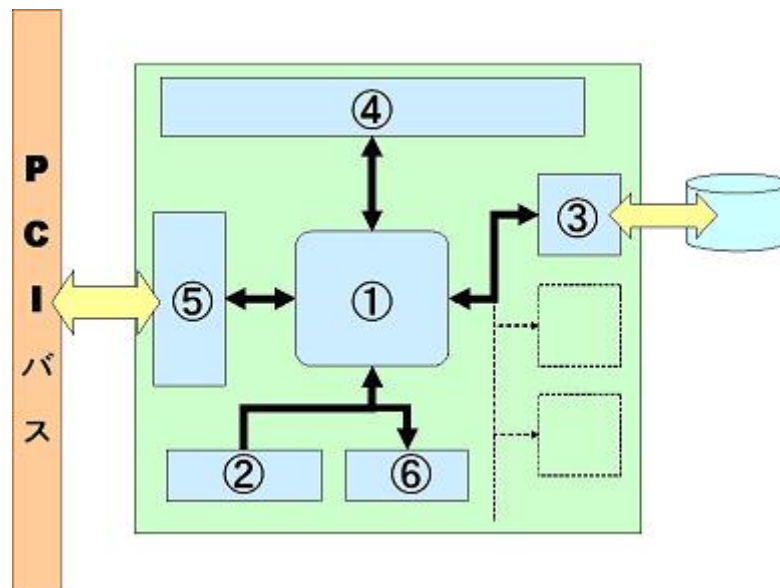
また RAID 処理専用のマイクロプロセッサを搭載したインテリジェントタイプと、ほとんどの RAID 処理を本体装置の CPU 上でデバイスドライバが実行するノンインテリジェントタイプがあり、下記のような特長があります。

- ・ インテリジェントタイプ
 - 高信頼性の RAID5、RAID6 もサポートしています。
 - ほとんどの RAID 処理を専用のプロセッサで実行するため、本体装置の CPU やメインメモリ等のリソースに与える影響は小さくなります。
 - RAID のシステム構成はハードディスクドライブにのみ記憶されます。そのため、故障したハードディスクドライブの状態によっては、RAID システムの回復に影響を及ぼす可能性があります。
 - RAID のログ情報は本体装置に保存されます。これにより、障害発生時からの復旧や障害発生原因の解析を容易に行う事ができます。
- ・ ノンインテリジェントタイプ
 - 高信頼性の RAID5、RAID6 はサポートしていません。
 - オンボード RAID による本体装置の CPU 使用率は約 2%~12%です。
 - RAID のシステム構成はハードディスクドライブにのみ記憶されます。そのため、故障したハードディスクドライブの状態によっては、RAID システムの回復に影響を及ぼす可能性があります。
 - RAID のログ情報は本体装置に保存されます。これにより、障害発生時からの復旧や障害発生原因の解析を容易に行う事ができます。
 - RAID 処理はデバイスドライバで実行するため、本体装置の電源 ON からドライバがロードされるまでの間に発生した障害やイベントのログが保存されない場合があります。

高信頼性、高耐障害性および高冗長性を必要されるシステムや、Linux OS を使用する場合は、インテリジェントタイプの RAID コントローラーを選択することを推奨します。

1.2.1 PCI カードタイプ

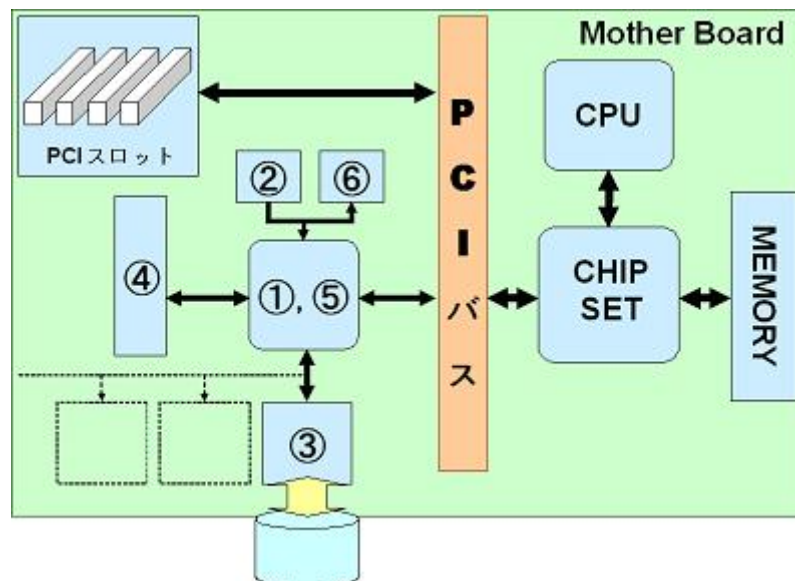
PCI カードタイプの RAID コントローラーは、RAID 処理専用マイクロプロセッサを搭載しており、ほとんどの RAID 処理を RAID コントローラー単体でおこなうためシステムのパフォーマンスに影響を与えません。



①	マイクロプロセッサ(MPU)	サポートする RAID レベルに必要な処理を全て担う専用プロセッサ
②	Flash ROM	マイクロプロセッサを制御するソフトウェアを格納するメモリ
③	インタフェースコントローラー	RAID コントローラーに接続する各種ハードディスクドライブに対応したインタフェースを制御するコントローラー
④	メモリ	パリティ処理やハードディスクドライブへの読み出し、書き込みに使用するキャッシュメモリ
⑤	PCIブリッジ回路	RAID コントローラーと PCI バスを接続するためのバスインタフェース
⑥	NvRAM	RAID システム構成、設定情報を記録するメモリ

1.2.2 メザニンカードタイプ

メザニンカードタイプの RAID コントローラーは、PCI スロットを介さずにマザーボード上に実装しています。ほとんどの RAID 処理を専用マイクロプロセッサにて行うため、システムのパフォーマンスに影響を与えません。



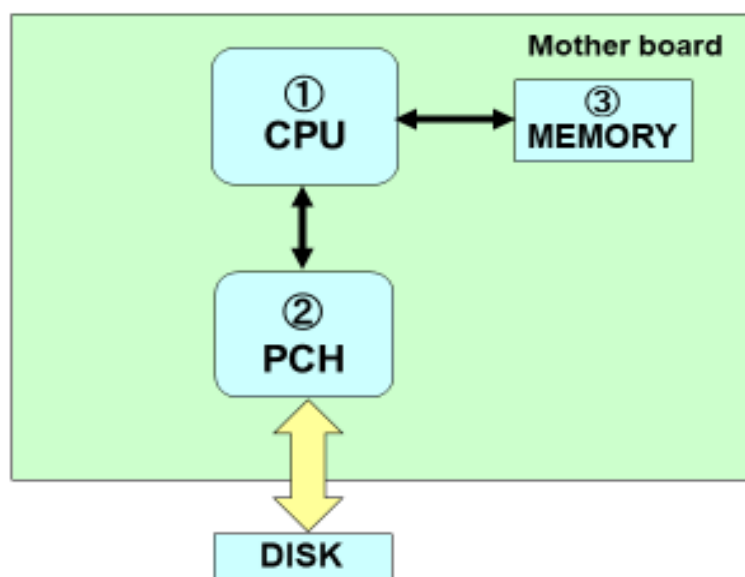
①	マイクロプロセッサ(MPU)	サポートする RAID レベルに必要な処理を全て担う専用プロセッサ
②	Flash ROM	マイクロプロセッサを制御するソフトウェアを格納するメモリ
③	インタフェースコントローラー	RAID コントローラーに接続する各種ハードディスクドライブに対応したインタフェースを制御するコントローラー
④	メモリ	パリティ処理やハードディスクドライブへの読み出し、書き込みに使用するキャッシュメモリ
⑤	PCI ブリッジ回路	RAID コントローラーと PCI バスを接続するためのバスインタフェース
⑥	NvRAM	RAID システム構成、設定情報を記録するメモリ

1.2.3 オンボードタイプ

オンボードタイプの RAID コントローラは、インタフェースコントローラも含めた全てのモジュールをマザーボード上に実装しています。RAID システムを安価に構築でき、ハードウェアリソースも削減できるメリットがあります。

① ノンインテリジェントタイプ

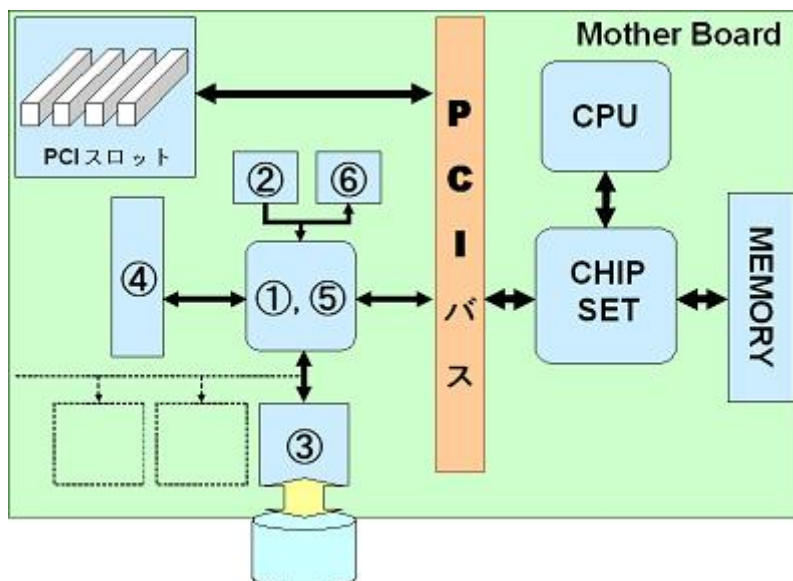
マイクロプロセッサをマザーボードに実装していないタイプです。ほとんどの RAID 処理を本体装置の CPU のデバイスドライバで行います。ノンインテリジェント カードタイプと同様、本体装置 CPU の使用状況により性能が上下します。



①	CPU	
②	PCH (プラットフォーム・コントローラー・ ハブ)	いくつかのデータパスを管理し、Intel 製 CPU の機能のサポートする役割。ノースブリッジ と サウスブリッジ を用いていたインテル・ハブ・アーキテクチャ の後続
③	MEMORY	

②インテリジェントタイプ

RAID 処理をおこなう専用マイクロプロセッサをマザーボードに実装しているタイプです。
ほとんどの RAID 処理を専用マイクロプロセッサにて行うため、システムのパフォーマンスに影響を与えません。



①	マイクロプロセッサ(MPU)	サポートする RAID レベルに必要な処理を全て担う専用プロセッサ
②	Flash ROM	マイクロプロセッサを制御するソフトウェアを格納するメモリ
③	インタフェースコントローラ	RAID コントローラに接続する各種ハードディスクドライブに対応したインタフェースを制御するコントローラ
④	メモリ	パリティ処理やハードディスクドライブへの読み出し、書き込みに使用するキャッシュメモリ
⑤	PCI ブリッジ回路	RAID コントローラと PCI バスを接続するためのバスインタフェース
⑥	NvRAM	RAID システム構成、設定情報を記録するメモリ

1.3 PCI 規格

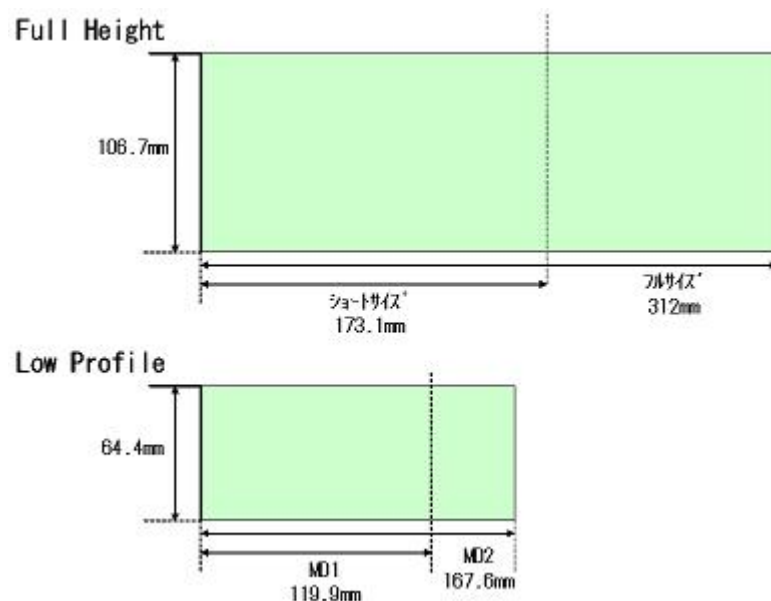
PCIとはPCI SIG(PCI Special Interest Group)により策定されているバスアーキテクチャのことです。従来のPC互換機にて最も多く使用されてきた拡張バスISA(Industry Standard Architecture)と比べ、機能面や性能面でも優り、現時点での業界標準となっています。

1.3.1 PCI Express

2002年にPCI-SIGによって策定された、PCIバスに代わるパソコン、サーバ向けシリアル転送インタフェースであり、3GIOを標準規格化したものです。PCIバスはパラレル転送方式を使用しており、PCI Expressとの間に物理レベルでの互換性はありませんが、通信プロトコルなどは共通のものが使われています。最小構成の伝送路(レーン)は片方向2.5Gbps(双方向5.0Gbps)の全二重通信が可能で、8ビットのデータを送るのにクロック信号など2ビットを追加した10ビットを費やすため、実効データ転送レートは片方向250MB/s(双方向500MB/s)です。また、片方向500MB/s(双方向1Gb/s)に向上したPCI Express 2.0、片方向1Gb/s(双方向2Gb/s)に向上したPCI Express 3.0も策定されています。実際のPCI Expressポートはこのレーンを複数束ねた構成になっていることが多く、1レーンで構成されたPCI Expressポートを「PCI Express x1」、2レーンのポートを「x2」といった具合に呼称します。

1.3.2 PCI ボードのサイズ

PCIボードとはPCI規格に適合した拡張ボードのことを指します。PCIボードには物理的寸法においてもボード幅およびボード長にさまざまな種類があります。



1.4 物理デバイスのインタフェース

物理デバイスの種類にはハードディスクドライブ(HDD)とソリッドステートドライブ(SSD)があり、HDD は磁性体が塗布されたアルミ合金やガラスの円盤(磁気ディスク)を複数枚かさね磁気的にデータを入出力する記憶装置、SSD は不揮発性のフラッシュメモリを用いた記憶装置です。これらを本体装置や RAID コントローラーと接続するインタフェース規格として、SAS(Serial Attached SCSI)規格、および SATA(Serial ATA)規格があります。

1.4.1 SAS 規格

SCSI をシリアル化したのが SAS です。SAS のデータ転送速度は 300MB/s(3Gb/s)で Ultra320 SCSI のバス当たり 320MB/s より遅いように思えますが、SAS はポイント・ツー・ポイント接続なので、1 台のデバイスが 300MB/s を占有できます。それを考慮すれば、Ultra320 SCSI より高速と言えます。SAS は拡張性も非常に高く、SAS のホスト・コントローラーとエンド・デバイス(例えばハードディスクドライブ)は、中継デバイスとなる「SAS エキスパンダ」を通じて接続できます。小規模な構成ならホスト・コントローラーを中心としたスター型、大規模構成なら複数の SAS エキスパンダを利用したツリー型トポロジを採れます。デバイス間の距離は最長 8m(外部ケーブル)で、接続可能なエンド・デバイス数は、規格上 1 万 6384 台までとなっています。また、規格の拡張にともない、最大転送速度が 1200MB/s(12Gb/s, SAS3.0)と性能面が向上しています。

1.4.2 SATA 規格

IDE をシリアル化したのが SATA です。IDE とくらべ SATA はケーブルやコネクタなどの物理的な仕様が大きく変更されています。ケーブルは 7 芯のシリアル伝送用のケーブルに変更され、規格が保証する線長は IDE 規格の 45cm であったのに対し、SATA では 100cm まで保証されています。信号線の接続形態としては IDE 規格では 1 本の信号線に対し 2 台までの IDE 機器を接続する事が可能でしたが、SATA では 1 本の信号線に対し 1 台しか接続できません。その分電気的な特性は向上しており、最大転送速度は 150MB/s(1.5Gb/s)と IDE 規格よりも速く、後継の Revision では最大転送速度が 600MB/s(6Gb/s)となり、性能面が向上しています。

1.4.3 SAS/SATA 比較

特徴	SAS	SATA
Scalability (拡張性)	1 対 1 接続のため電気特性が高い。 (ノイズに強い)	
	1ch につき最大 1 台まで接続可能※1	1ch につき最大 1 台まで接続可能
Performance (能力)	1ch あたり最大転送速度 12Gb/s	1ch あたり最大転送速度 6Gb/s
	回転数 10,000～15,000rpm	回転数 5,400～7,200rpm
	複数同時処理能力が高い (コマンドキューイング機能あり)	複数同時処理能力が高い (コマンドキューイング機能がある製品に限り)
Reliability (信頼性)	リアサイン機能等の異常時のリカバリ処理能力が高く エラー発生時のステータス情報が豊富。	リカバリ処理能力が低く、 エラー発生時のステータス情報が乏しい。
	メディア記録密度が低い分、塵等の影響に強く Head マージンも高くなるためエラーが少ない。	
Maintainability (保守性)	活栓挿抜可能な標準インタフェースを HDD に持つため、システム稼動中に HDD 交換が可能。	
Cost (費用)	メディア記録密度が SATA に比べ低い GB あたりの単価が高い	メディア記録密度が SAS に比べ高い GB あたりの単価が高い

※1: SAS エキスパンダを利用することで、1ch に複数台接続可能

1.4.4 SAS/SATA の奨励用途

物理デバイスのインタフェース	奨励用途	主な理由
SAS	データベースサーバなどのハイエンド環境	高い性能、および信頼性を有する
SATA	ワークグループサーバなどの SOHO 環境	価格メリットを活かし、用途に応じサーバを使い分けるシステム展開が可能

1.4.5 セクタサイズ

HDD のセクタサイズには、媒体上に書き込まれるデータ単位、I/F 上でやり取りされるデータ通信単位により、512 native(512n)、4K native(4Kn)、512 emulation(512e)の3種類に分けられます。本タイプの RAID コントローラーは4Knをサポートしていません。

		媒体上の記録単位(Byte)	
		512	4096
データ通信単位 (Byte)	512	512n	512e
	4096	—	4Kn

4Kn/512e はより大容量の HDD を実現するための技術で、媒体上に記録するデータの単位を従来の 512n HDD より大きくしたものです。4Kn と 512e の違いは、ホスト・コントローラーとのデータ通信の単位です。4Kn は対応するホスト・コントローラーに接続することで、媒体上に記録されるデータ単位と同じ 4096Byte 単位で通信を行います。512e は従来のホスト・コントローラーに接続可能とするため、媒体上に記録されたデータとは異なる 512Byte 単位に分割してデータ通信を行います。全ての通信が分割されるため、512e では見かけ上のメディアエラーの発生数(4Kn の 8 倍)、メディアエラーによる HDD 交換基準(160 回)は 4Kn と比べて多くなりますが、品質/性能への影響の程度は 4Kn と変わりません。

第2章 機能編

本章では RAID コントローラーの機能を説明します。
各製品のサポート RAID レベル、キャッシュ容量、ポート数の情報を下記に記載します。

機能対応表

	対応 RAID レベル									キャッシュ	機能 HBA パススルー	ス ペ ア デ ィ ス ク	RAID 拡張	ポ ー ト 数
	RAID 0	RAID 1	RAID1(ADM)	RAID 5	RAID6	RAID10	RAID10(ADM)	RAID50	RAID60					
N8103-189	○	○	×	○	×	○	×	×	×	なし	○	○	○	内蔵8ポート
N8103-190	○	○	○	○	○	○	○	○	○	2GB	○	○	○	内蔵8ポート
N8103-191	○	○	○	○	○	○	○	○	○	4GB	○	○	○	内蔵16ポート
N8103-192	○	○	×	○	×	○	×	×	×	なし	○	○	○	内蔵8ポート
N8103-193	○	○	○	○	○	○	○	○	○	2GB	○	○	○	内蔵8ポート
N8103-194	○	○	○	○	○	○	○	○	○	4GB	○	○	○	内蔵16ポート
N8103-195	○	○	×	○	×	○	×	×	×	なし	○	○	○	内蔵8ポート
N8103-201	○	○	○	○	○	○	○	○	○	2GB	○	○	○	内蔵8ポート
N8103-196	○	○	○	○	○	○	○	○	○	4GB	○	○	○	外部8ポート
N8103-237	○	○	○	○	○	○	○	○	○	4GB	○	○	○	内蔵16ポート
N8103-238	○	○	○	○	○	○	○	○	○	8GB	○	○	○	内蔵32ポート
オンボード RAID (SmartArrayS100i)	○	○	×	×	×	○	×	×	×	なし	×	○	×	本体装置に依存

○・・・機能あり

×・・・機能なし

2.1 RAID システム構築機能

2.1.1 ディスクアレイ(パック)

ディスクアレイ(パック)とは複数の物理デバイスのグループを表し、論理ドライブを設定するための基となります。
設定可能なディスクアレイ数は、本体装置やディスク増設ユニットに搭載した物理デバイスの台数、ディスクアレイの種類(RAID レベル)、および RAID コントローラーの最大作成可能ディスクアレイ数により異なります。*1

2.1.2 ディスクアレイ(パック)の構成ルール

ディスクアレイ(パック)は以下のルールに則り構成する必要があります。

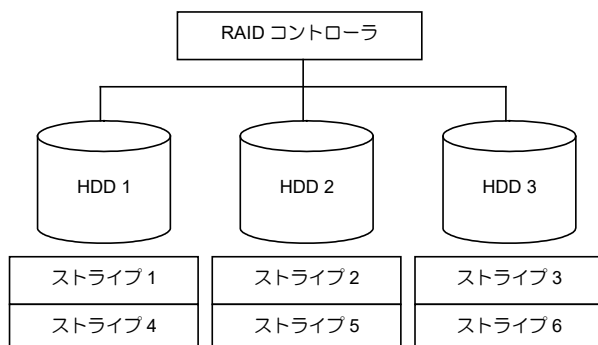
[ルール]

- 同一 RAID コントローラー配下の物理デバイスを使用して、(RAID レベルを問わず)複数のディスクアレイを組むことが可能です
- 同一 RAID コントローラー配下の物理デバイスを使用して、複数のディスクアレイを組むことが可能です
- RAID コントローラーをまたいだ物理デバイスを使用して、ディスクアレイを組むことはできません。

2.1.3 RAID の種類(RAID レベル)

RAID0 (ストライピング)

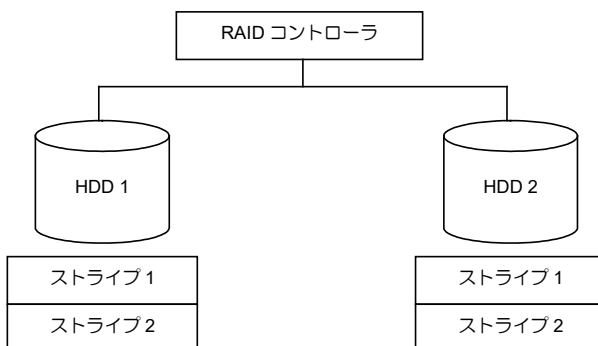
複数台の物理デバイスを単一ドライブに見立て、アクセスを分散する事で高速化、大容量化を実現します。



RAID0 の特徴	
冗長性	無し
特徴	全物理デバイス容量をデータディスクとして使用可能
	RAID レベルの中で最も高速
	冗長構造ではないため物理デバイスが故障(Dead)するとデータをロストしてしまう
使用に適した AP	クリティカルでないデータに対して高い性能を必要とする AP
ドライブ数	1 台以上

RAID1 (ミラーリング)

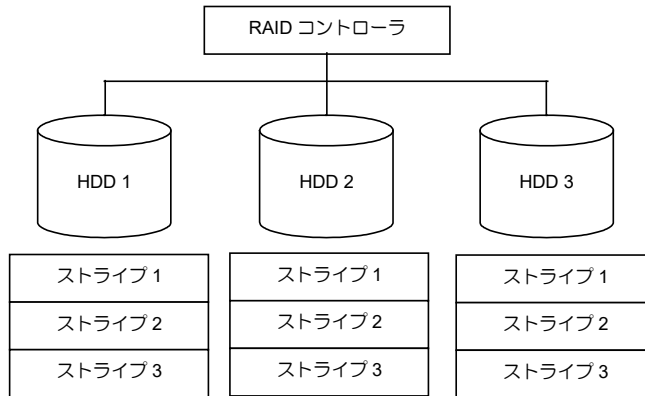
2 台 1 組の物理デバイスに対し常に同じデータを格納する事でデータを二重化し高信頼性を実現します。



RAID1 の特徴	
冗長性	有り
特徴	1 台の物理デバイスが故障(Dead)しても、もう片方の複製物理デバイスより復旧をおこなう
	2 台の物理デバイスのみで冗長性のある RAID システムを構築できるため、必要な総コストは最も低くなる
	データを書き込める容量は物理デバイスの総容量の 1/2 になる
使用に適した AP	論理ドライブ、重要なファイルを格納するドライブ
ドライブ数	2 台

RAID1 (ADM)

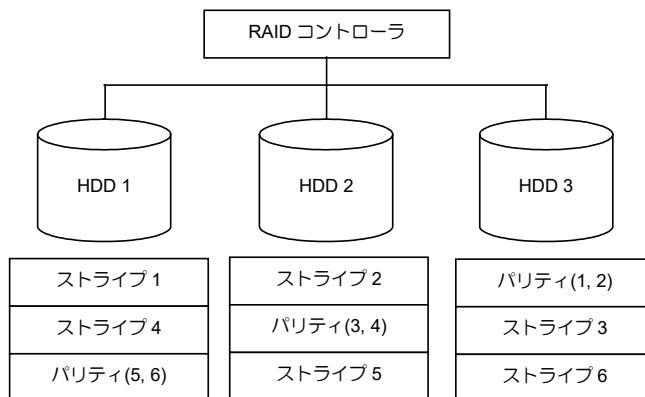
3 台 1 組の物理デバイスに対し常に同じデータを格納する事でデータを三重化し RAID1 よりも高い信頼性を実現します。



RAID1(ADM)の特徴	
冗長性	有り
特徴	2 台の物理デバイスが故障(Dead)しても、残り 1 台の複製物理デバイスより復旧をおこなう
	3 台の物理デバイスで冗長性のある RAID システムを構築する。RAID1 よりもコストがかかる。
	データを書き込める容量は物理デバイスの総容量の 1/3 になる
使用に適した AP	論理ドライブ、重要なファイルを格納するドライブ
ドライブ数	3 台

RAID5 (パリティ付きストライピング)

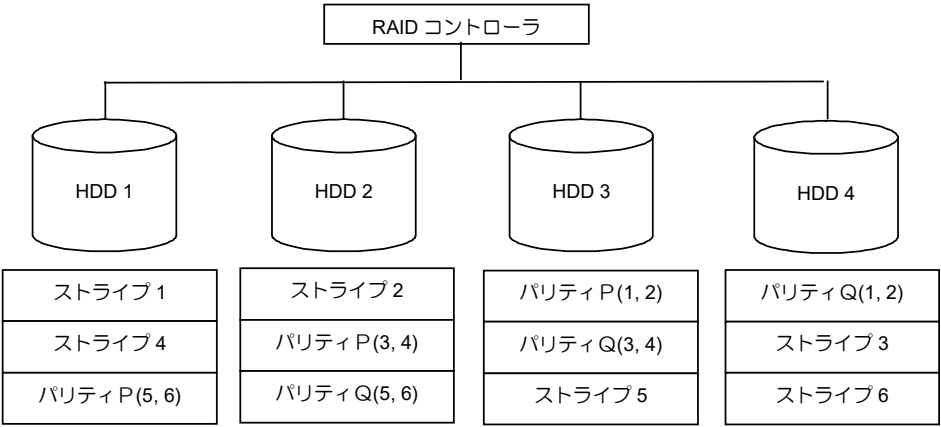
複数台の物理デバイスを単一ドライブに見立て、アクセスを分散します。また、保存するデータのパリティを生成し各物理デバイスに保存します。これにより高速化、大容量化および高信頼性を実現します。



RAID5 の特徴	
冗長性	有り
特徴	1 台の物理デバイスが故障(Dead)してもデータを保護することができる
	大きなファイルのシーケンシャル読み出しが高速である
	データ以外にパリティを物理デバイスに格納するため物理デバイス総容量の 66%～がデータを格納できる領域となる
	パリティを生成する時間がかかるため、書き込み性能は高くない
使用に適した AP	重要なデータを大量に扱い、リード性能が要求される AP
ドライブ数	3 台以上

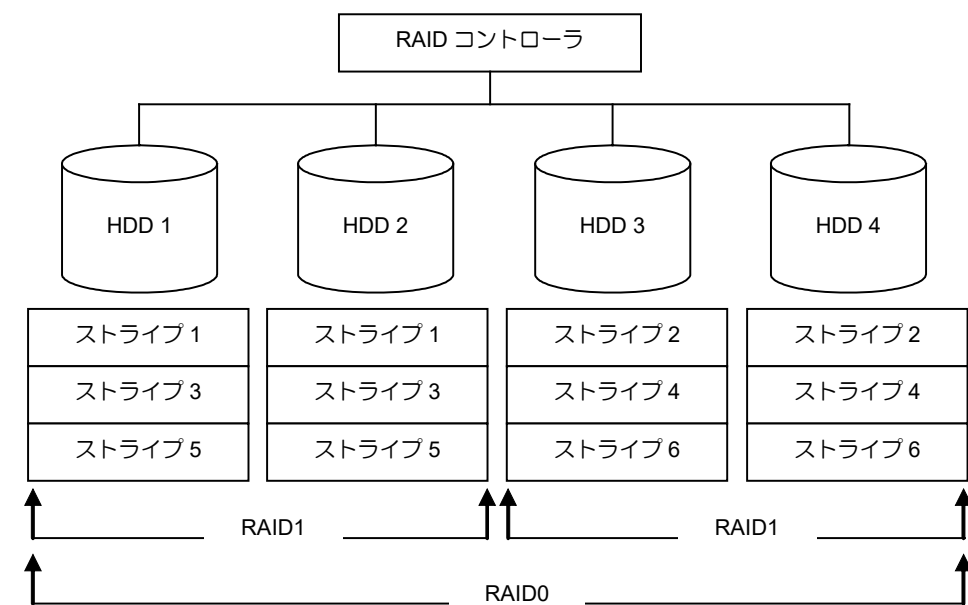
RAID6 (二重化パリティ付きストライピング)

複数台の物理デバイスを単一ドライブに見立て、アクセスを分散します。また、保存するデータのパリティを生成し、各物理デバイスに二重化して保存します。これにより高速化、大容量化および高信頼性を実現します。



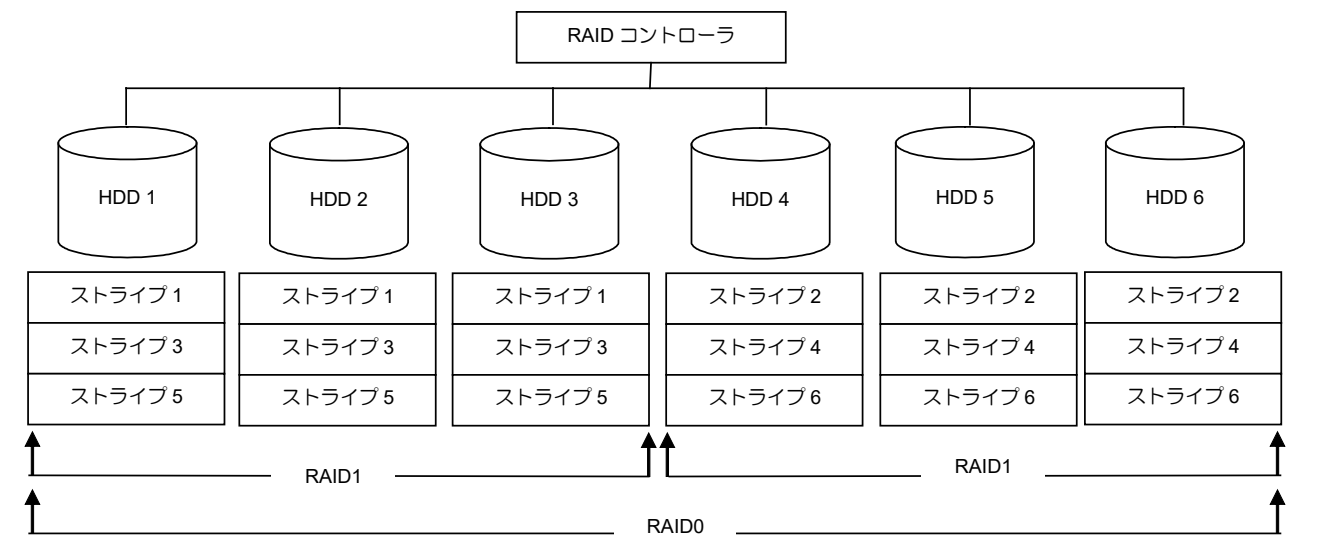
RAID6 の特徴	
冗長性	有り
特徴	2 台の物理デバイスが故障(Dead)してもデータを保護することができる
	大きなファイルのシーケンシャル読み出しが高速である
	データ以外にパリティを物理デバイスに格納するため物理デバイス総容量の 33.3%～がデータを格納できる領域となる
	パリティを生成する時間がかかるため、書き込み性能は高くない
使用に適した AP	重要なデータを大量に扱い、リード性能が要求される AP
ドライブ数	4 台以上

RAID10



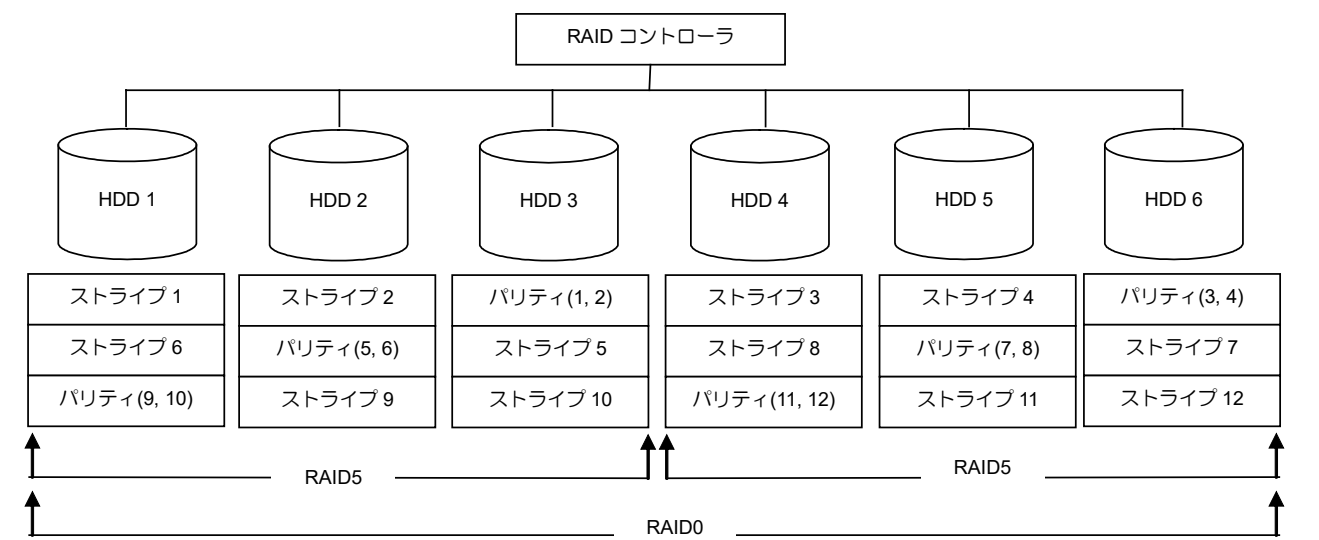
RAID10 の特徴	
冗長性	有り
特徴	1～2 台の物理デバイスが故障(Dead)してもデータを保護することができる(物理デバイス 2 台故障(Dead)の場合は組み合わせによる)
	書き込み性能は RAID1 を多少上回る
	データを書き込める容量は物理デバイスの総容量の 1/2 になる
使用に適した AP	論理ドライブ、重要なファイルを格納するドライブ
ドライブ数	4 台以上

RAID10 (ADM)



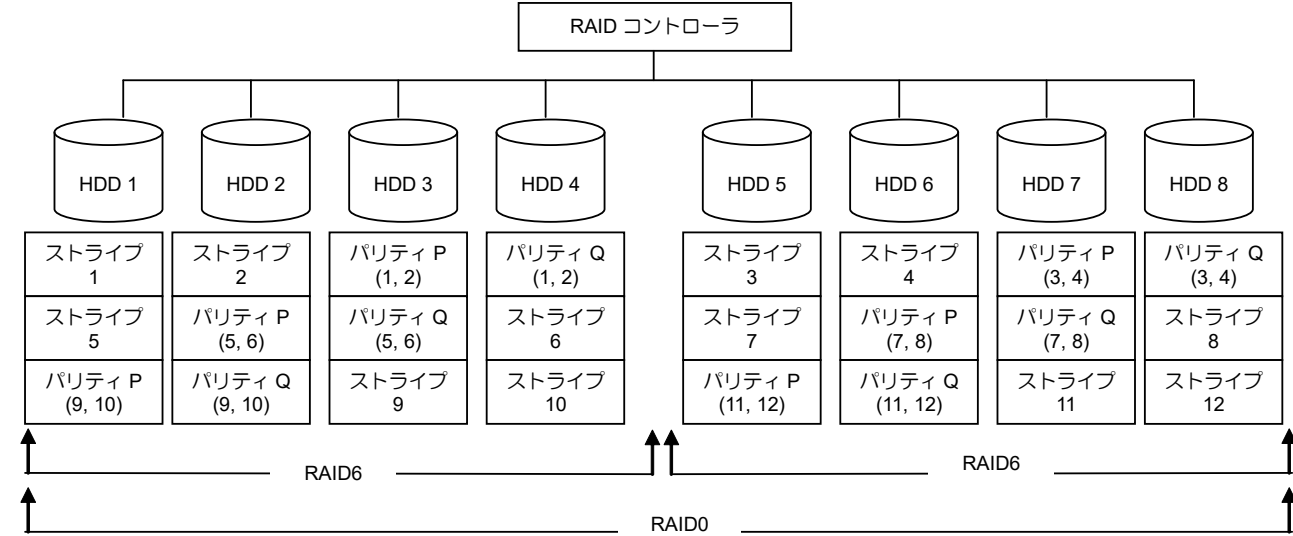
RAID10(ADM)の特徴	
冗長性	有り
特徴	1～4 台の物理デバイスが故障(Dead)してもデータを保護することができる(物理デバイス 3～4 台故障(Dead)の場合は組み合わせによる)
	書き込み性能は RAID1(ADM)を多少上回る
	データを書き込める容量は物理デバイスの総容量の 1/3 になる
使 用 に 適 した AP	論理ドライブ、重要なファイルを格納するドライブ
ドライブ数	6 台以上

RAID50



RAID50 の特徴	
冗長性	有り
特徴	1～2 台の物理デバイスが故障(Dead)してもデータを保護することができる(物理デバイス 2 台故障(Dead)の場合は組み合わせによる)
	書き込み性能は RAID5 を多少上回る
	大きなファイルのシーケンシャル読み出しが高速である
	データ以外にパリティを物理デバイスに格納するため物理デバイス総容量の 66%～がデータを格納できる領域となる
使用に適した AP	重要なデータを大量に扱い、リード性能が要求される AP
ドライブ数	6 台以上

RAID60



RAID60 の特徴	
冗長性	有り
特徴	2～4 台の物理デバイスが故障(Dead)してもデータを保護することができる(物理デバイス 4 台故障(Dead)の場合は組み合わせによる)
	大きなファイルのシーケンシャル読み出しが高速である
	データ以外にパリティを物理デバイスに格納するため物理デバイス総容量の 33.3%～がデータを格納できる領域となる
	パリティを生成する時間がかかるため、書き込み性能は高くない
使用に適した AP	重要なデータを大量に扱い、リード性能が要求される AP
ドライブ数	8 台以上

2.2 初期化

初期化機能とは論理ドライブを構築している物理デバイスに対し、初期化処理を行う機能です。

RAID コントローラーは複数の物理デバイスを論理的に結合することで論理ドライブを構築することができます。しかし、論理ドライブを構築している物理デバイスのすべてが新品であるなどの場合、物理デバイス内のデータが消去されているとは限りません。そのため、初期化機能を使用して論理ドライブを構築している物理デバイスに対し初期化処理を行います。初期化はデフォルトと迅速の2種類があります。

2.2.1 初期化方法

① デフォルト

パリティを使用する RAID5,6,50,60 において、パリティブロックを有効な値に初期化する機能です。

表面スキャン分析と同等の処理を行います。

※パリティの初期化プロセスが完了する前に、オペレーティングシステムから論理ドライブを認識し、使用することができません。

② 迅速

パリティを使用する RAID5,6,50,60 において、パリティブロックを有効な値に初期化する機能です。

データブロックとパリティブロックの両方を初期化します。

※パリティの初期化プロセスが完了するまで、オペレーティングシステムから論理ドライブを認識し、使用することができません。

2.3 再構築(リビルド)

再構築(リビルド)は、論理ドライブを構築している物理デバイスが故障(Dead)した場合、障害が発生した物理デバイスを正常な物理デバイスと交換することで、元の正常な論理ドライブを再構築する機能です。

RAID1 や RAID5、RAID6、RAID10、RAID50、RAID60 といった冗長性のある論理ドライブに対して実行することができます。

2.3.1 ホットスペア(スタンバイ)リビルド

冗長性のある RAID システムにて、ホットスペアをあらかじめ RAID システムに組み込み、物理デバイスの障害発生時に自動的にホットスペアを用いて実行されるリビルドをホットスペア(スタンバイ)リビルドといいます。

2.3.2 ホットスワップリビルド

冗長性のある RAID システムにて、システム稼動中でも電源を落とすことなく、障害が発生した物理デバイスを交換する、この機能をホットスワップと呼びます。そしてホットスワップにて交換された物理デバイスに対して自動的に実行されるリビルドをホットスワップリビルドといいます。

2.3.3 リビルド時間目安

完了までに必要な時間目安については、付録 C を参照してください。

2.3.4 スペアドライブ

スペアドライブの管理モードには“専用”と“自動交換ドライブ”があります。

- ・スペアの管理モード: 専用

1 つの専用スペアドライブは、同一 RAID コントローラー配下の複数のアレイを対象として設定することができます。障害が発生した物理デバイスを交換すると、コピーバックが動作します。このため、スペアドライブのスロットは変わりません。

- ・スペアの管理モード: 自動交換ドライブ

自動交換ドライブが動作した場合、障害が発生した物理デバイスの代わりに対象のアレイに組み込まれます。障害が発生した物理デバイスを交換すると、そのまま自動交換ドライブとなり、コピーバックは動作しません。このため、スペアドライブのスロットが交換した物理デバイスのスロットに変わります。

スペアのアクティベーションモードには“予測スペアのアクティベーション”と“障害スペアのアクティベーション”があります。

- ・予測スペアのアクティベーション

予測スペアアクティベーションは、スペア対象のアレイの物理デバイスが障害予測 (S.M.A.R.T.) ステータスを報告した場合に、再構築を動作させる機能です。アレイがオンラインの状態、S.M.A.R.T.の物理デバイスからスペア対象の物理デバイスにデータをコピーします。

データのコピー完了後、S.M.A.R.T.の物理デバイスは、縮退し、交換待ちの状態になります。交換後は、スペア対象の物理デバイスから交換した物理デバイスにデータを書き戻し、元のスロットの状態に戻します。



本機能には、以下の特徴があります。:

バッドブロックを修復することができます。

RAID0 を含む、全ての RAID レベルで使用可能です。

- ・障害スペアのアクティベーション

障害スペアのアクティベーションは、冗長性のあるアレイを構成する物理デバイスが縮退した場合に、動作します。

リビルドの完了までに必要な時間目安については、付録 C を参照してください。

2.4 表面スキャン

表面スキャンは自動的にバックグラウンドで動作し、物理デバイスに不良が発生した際にデータの修復を行います。表面スキャンには以下の特徴があります。

- ・冗長性のある論理ドライブを構成する物理デバイスに対して不良セクタのチェックを行います。
- ・RAID5 もしくは RAID6 の構成に対してはパリティデータの整合性のチェックを行います。

表面スキャンが動作していない物理デバイスに対して、表面スキャンの優先順位を無効、高、アイドルに設定することができます。

- ・無効: 無効にすることで表面スキャンの I/O が完了するまでのレイテンシの影響を減らすことができますが、データ損失につながる不良セクタが発生する可能性が高まります。
- ・高: 表面スキャンの設定を高にすることでデータが損失する前に不良ブロックを検出する可能性が高まります。
- ・アイドル: 表面スキャンをアイドルにし、対応する遅延時間を設定することで、レイテンシの影響を低減させ、アイドル時の不良ブロックのスキャンを行うことができます。

並列表面スキャン数は表面スキャンが並列に実行できる数を設定します。コントローラー上に一つ以上の論理ドライブが存在している場合に使用されます。この設定はコントローラーが複数の論理ドライブの不良ブロックを検出し、容量の大きなドライブを使用している論理ドライブの検出にかかる時間を低減させることができます。

2.5 電力モード

次の 3 つの利用可能な電力モードがあります。

- ・最大パフォーマンス
- ・バランス
- ・省電力

<最大パフォーマンス(デフォルト)>

これはデフォルト設定です。すべての設定は、最大のパフォーマンスに基づいて選択されます。パフォーマンスに影響する電力節約オプションは無効です。

<バランス>

パフォーマンスへの影響を最小限に抑えて電力を節約するにはこの設定を使用します。キューの項目数が多い場合に、この設定がスループットに与える影響は 10% 以下です。

キューの項目数が少ない、または I/O が頻繁ではない場合、パフォーマンスへの影響は大きくなる場合があります。このコマンドは、通常、ハードディスクドライブのみを使用する環境で役立ち、SSD 使用時にはお勧めしません。

設定は、ドライブの数やタイプ、RAID レベル、ストレージのトポロジなど、ユーザーの構成に基づきます。構成を大幅に変更すると、最適な設定を選択するために再起動が必要となる場合があります。設定を変更するために再起動が必要な場合、SSA は警告を生成します。

<省電力>

システムパフォーマンスにこだわらずにこの設定を選択すれば、最大の電力の節約が実現されます。ごく一部のアプリケーションにはこの設定をおすすめしていますが、ほとんどのお客様に適切な設定ではありません。ほとんどのアプリケーションにおいて大幅なパフォーマンスの低下が生じます。

2.6 キャッシュ機能 Flash-Backed write cache(FBWC)

本製品は、ライト時にRAIDコントローラーのキャッシュメモリを使い、ライト性能を大幅に向上させることができます。また、停電が発生した場合にキャッシュメモリ内のデータをバックアップすることができます。

キャッシュ設定の変更では、読み取りと書き込みに割り当てるメモリ容量を設定します。アプリケーションにはそれぞれ最適な設定があります。RAIDコントローラー上で論理ドライブが作成されていて、キャッシュを持っている場合にのみ設定を変更することができます。

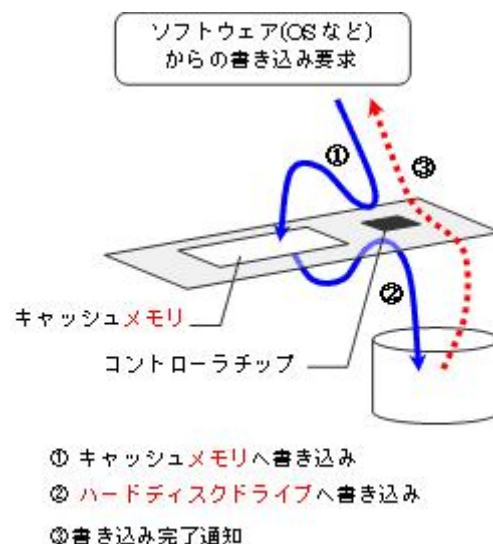
初期値は書き込み90%、読み取り10%です。データが連続する場合や、新しい書き込みデータから読み取りが多く予想される場合は、読み取りの割合を高く設定するとより効果が得られます。

2.6.1 無効

どのソフトウェアから書き込み要求がきた場合に、RAIDコントローラー上のキャッシュメモリと物理デバイスの両方に書き込みを行う方式。

ソフトウェアは、物理デバイスへの書き込み処理が終了するのを待ってから次の処理に移るため、一般的にアクセス性能は劣ります。

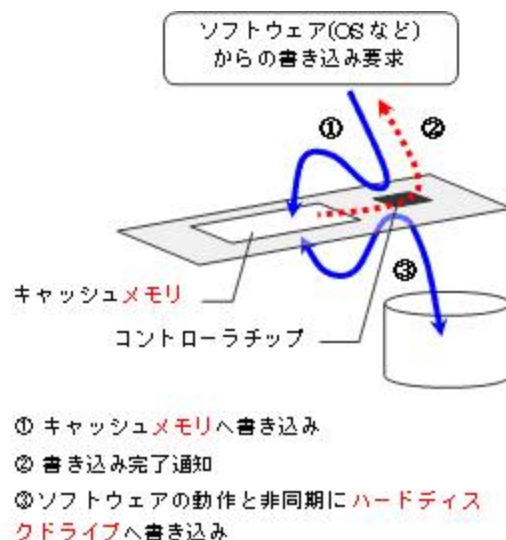
しかし、ソフトウェアからの書き込み要求が即時に物理デバイスに反映されるため、電源瞬断などの不慮の事故が発生してもデータを損失する危険性が少ないという利点があります。



2.6.2 有効

OSなどのソフトウェアから書き込み要求がきた場合に、RAIDコントローラー上のキャッシュメモリへのみ書き込みを行い、物理デバイスへの書き込みはキャッシュメモリ上のデータを元にRAIDコントローラーが非同期に行う方式。

キャッシュメモリにデータが書き込まれた時点でソフトウェア側に完了通知が発行されるため、物理デバイスへの書き込み処理が完了するのを待たずにソフトウェア側は次の処理を継続することができます。一般的にアクセス性能が向上しますが、電源瞬断などの不慮の事故が発生した際にキャッシュメモリの内容が物理デバイス上に反映されない場合があり、データ損失の危険性があります。



2.6.3 バッテリー

本体装置にバッテリーを接続することで、サーバに電源が供給されていない間(「キャッシュデータ保持時間」の範囲で)キャッシュ上にデータを保持します。この機能により、キャッシュ有効で運用しているシステムにおいて、電源瞬断などの不慮の事故によるデータ損失を防ぐことができます。

バッテリーが寿命を迎えた場合、キャッシュモードが自動的に切り替わり、アクセス性能が低下した状態で運用され続けるリスクがあります。製品寿命(3年)を迎えたバッテリーは速やかに交換してください。

※キャッシュデータは不揮発性メモリに保持するため、長期の保持が可能です。

※本バッテリーは1個搭載ですべてのRAIDコントローラーに電力供給可能です。

2.6.4 フラッシュバックアップユニット

本体装置にフラッシュバックアップユニット(FBU)を接続することで、停電やサーバ障害時にFBUの電力を使用し、キャッシュデータをフラッシュメモリに転送します。フラッシュメモリは不揮発性メモリのためメモリ保持に電力が不要です。この機能により、キャッシュ有効で運用しているシステムにおいて、電源瞬断などの不慮の事故によるデータ損失を防ぐことができます。

フラッシュバックアップユニットではバッテリーを使わない方式を採用しているため、定期交換の必要がありません。また、定期交換不要なため、システムのダウンタイムの削減にも繋がります。FBUは、バッテリーに比べて充電量が少ないため、複数枚のRAIDコントローラーを接続する場合にはバッテリーを選択してください。具体的な搭載条件はシステム構成ガイドを参照してください。

2.7 アレイ変換と論理ドライブの移行

2.7.1 アレイの拡張

割り当てられていない物理デバイスを使用して、作成済のアレイの容量を増やします。使用する割り当てられていないドライブは以下の条件を満たす必要があります。

※使用する物理デバイスは、既存の物理デバイスと同一容量、同一規格のものを使用してください。

2.7.2 アレイの移動

アレイの移動は、作成済のアレイを、同じ構成のアレイ、または、同一セットの物理デバイスに移動します。アレイの移動を使用するには、以下の条件を満たす必要があります。

- ・移動先のアレイを構成する物理デバイスの数が対象のアレイと同数であること
- ・移動先のアレイを構成する物理デバイスのインタフェース(例: SATA /SAS)が対象のアレイを構成する物理デバイスと同一であること
- ・移動先のアレイの容量が対象のアレイの容量より大きいこと

2.7.3 アレイの交換

アレイの交換は、作成済のアレイの全ての論理ドライブを、他のアレイに移行する機能です。アレイの交換を実行したアレイは削除され、割り当てられていないドライブとなります。アレイの交換を使用するには、以下の条件を満たす必要があります。

- ・交換先のアレイが交換元のアレイと同じ物理デバイス数で構成されている
- ・交換先のアレイ、および、交換元のアレイが共に OK のステータスである。(交換元のアレイ中の論理ドライブが全て OK のステータスである)
- ・交換先のアレイに、交換元のアレイに含まれるすべての論理ドライブを収納できるだけの、十分な容量がある

2.7.4 アレイの縮小

アレイの縮小は、作成済のアレイから物理デバイスを削除する機能です。アレイの縮小を使用するには、以下の条件を満たす必要があります。

- ・縮小後のアレイに、構成されているすべての論理ドライブを収納できるだけの十分な容量が必要です。
- ・アレイから物理デバイスを削除した結果、物理デバイスの数が既存の論理ドライブのフォールトトレランス(RAID レベル)をサポートできなくなる場合は削除できません。たとえば、4 台の物理デバイスと RAID5 論理ドライブを含むアレイがある場合、RAID5 では 3 台以上の物理デバイスが必要なので、削除できる物理デバイスの数は 1 台だけです。
- ・アレイに RAID10 論理ドライブが含まれる場合、削除する物理デバイスの数は偶数でなければなりません。
- ・アレイに複合タイプの RAID (RAID50 または RAID60) の論理ドライブが含まれる場合、削除する物理デバイスの数はパリティグループの数の倍数でなければなりません。たとえば、10 台の物理デバイスと RAID50 論理ドライブが含まれるアレイを縮小する場合、削除できる物理デバイスの数は 2 台または 4 台だけです。

2.7.5 アレイのミラー化

アレイのミラー化は、RAID0 のアレイを RAID1、または、RAID10 にミラーリングする機能です。アレイのミラー化を使用するには、以下の条件を満たす必要があります。

- ・アレイのミラー化を実行した後の物理デバイスの数が 2 台の場合は RAID1 となり、4 台以上の場合は RAID10 となります。

2.7.6 アレイの修復

アレイの修復は、障害が発生した物理デバイスを、別の正常な物理デバイスに交換する機能です。交換した後でも、元のアレイと論理ドライブの番号は影響を受けません。アレイの修復を使用するには、以下の条件を満たす必要があります。

- ・交換用物理デバイスと元の物理デバイスのインターフェイスタイプ(SAS、SATA など)が同じである。
- ・アレイ内の障害を起こした各物理デバイスを交換するために、アレイ上の最も小さい物理デバイスと同等以上のサイズの割り当てられていない物理デバイスが十分な台数が存在している。
- ・アレイ内に障害を起こした物理デバイスが 1 台以上ある。
- ・(スペアの再構築など)アレイの変換が行われていない。
- ・アレイの変換を実行できる動作中のキャッシュがある。

2.8 論理ドライブの移行

2.8.1 論理ドライブの拡大

作成済の論理ドライブの容量を拡大する機能です。拡大した領域は、OS より新たにパーティションを作成し、使用することができます。

2.8.2 RAID レベルの変更

作成済の論理ドライブの RAID レベルを変更する機能です。RAID レベルを変更した場合、容量が変化する場合があります。より冗長な RAID レベルに移行する場合、アレイに未使用の容量が必要になる可能性があります。移行したアレイで冗長性を保つための使用するデータが増えるため、この余分の容量が必要になります。

2.8.3 ストライプサイズの移行

作成済の論理ドライブのストライプサイズを変更する機能です。ストライプサイズを変更した際、容量が変化する場合があります。また、より大きなストライプサイズに移行する場合、アレイに未使用の容量が必要になる可能性があります。移行したアレイでより大きなデータストライプの一部が効率的に利用されていないために、この余分の容量が必要になります。

2.8.4 移行の優先度

通常の IO に対する、移行処理の優先度を設定します。本設定はアレイの拡張、論理ドライブの拡大、論理ドライブの移行、アレイの縮小に影響します。

- ・設定を高にすると、他のすべてのシステム動作よりも拡張が優先されます。
- ・設定を中にすると、拡張の時間が半分になり、残りの時間に通常のシステム動作が行われます。
- ・設定を低にすると、アレイの拡張よりも通常の動作が優先されます。

第3章 ハードウェア編

3.1 RAID コントローラー製品一覧

前述ように RAID は複数台の物理デバイスを用いて高速化、大容量化、高信頼性を実現するための技術です。RAID コントローラーとは RAID の持つ、優れた特性を発揮するための専用ハードウェアです。RAID コントローラーは、実装形態や、搭載されている集積回路の規模、接続可能なインタフェース等により大別する事ができます。下記に RAID コントローラーの製品一覧を示します。

① PCIカードタイプ製品一覧

N コード	インタフェース	コネクタ/ポート数	最大転送速度	PCI バス形式	タイプ
N8103-195	SAS/SATA	2 コネクタ/8port	SATA:6Gb/s SAS:12Gb/s (1port あたり)	PCI Express(x8)	Intelligent
N8103-196					
N8103-201					
N8103-238	SAS/SATA/ NVMe	4 コネクタ/32port	SATA:6Gb/s SAS:24Gb/s NVMe:16GT/s (1port あたり)	PCI Express(x16)	

② メザニンカードタイプ製品一覧

N コード	インタフェース	コネクタ/ポート数	最大転送速度	PCI バス形式	タイプ
N8103-189	SAS/SATA	2 コネクタ/8port	SATA:6Gb/s SAS:12Gb/s (1port あたり)	PCI Express(x8)	Intelligent
N8103-190		4 コネクタ/16port			
N8103-191		2 コネクタ/8port			
N8103-192		4 コネクタ/16port			
N8103-193					
N8103-194					
N8103-237	SAS/SATA/ NVMe	2 コネクタ/16port	SATA:6Gb/s SAS:24Gb/s NVMe:16GT/s (1port あたり)		

③ オンボード RAID 製品一覧

N コード	インタフェース	コネクタ/ポート数	最大転送速度	PCI バス形式	タイプ
オンボード RAID (SmartArrayS100i)	SATA	本体装置に依存	本体装置に依存	本体装置に依存	Non Intelligent

3.2 各 RAID コントローラーの仕様

各 RAID コントローラーの仕様については、付録 A を参照してください。

3.3 バックアップユニットの仕様

本体装置によって使用できるバッテリーが異なります。接続対象について詳しくはシステム構成ガイドを参照してください。
本バッテリーは 1 個搭載ですべての RAID コントローラーに電力供給可能です。

N コード	タイプ	データ 保持時間	充電 時間	定期リフレッシュ 推奨実行間隔 (リフレッシュ時間)	製品 寿命	制御ケーブル長
N8103-198	バッテリー	不揮発性メモリに保持するため、 長期の保持が可能	2 時間 (最長)	不要	3 年	145mm
N8103-203	バッテリー	不揮発性メモリに保持するため、 長期の保持が可能	2 時間 (最長)	不要	3 年	260mm
N8103-215	バッテリー	不揮発性メモリに保持するため、 長期の保持が可能	2 時間 (最長)	不要	5 年	230mm

3.4 フラッシュバックアップユニットの仕様

本体装置によって使用できるフラッシュバックアップユニット(FBU)が異なります。接続対象について詳しくはシステム構成ガイドを参照してください。

N コード	タイプ	データ 保持時間	充電 時間	定期リフレッシュ 推奨実行間隔 (リフレッシュ時間)	製品 寿命	制御ケーブル長
N8103-218	FBU	不揮発性メモリに保持するため、 長期の保持が可能	数分	不要	5 年	145mm
N8103-219	FBU	不揮発性メモリに保持するため、 長期の保持が可能	数分	不要	5 年	260mm

3.5 物理デバイス選定における確認事項

- ・ RAID システムに使用する物理デバイスは、新しい物理デバイスあるいは完全にフォーマットされた物理デバイスを使用してください。
- ・ ディスクアレイを構成するときは、同容量/同回転数/同規格の物理デバイスにて構成してください。
- ・ 容量の異なる物理デバイスを同一ディスクアレイに混在させた場合、ディスクアレイ内の物理デバイス 1 台あたりの容量は最も小さい物理デバイス容量に揃えられます。
- ・ サポートする本体装置、搭載ルールについては、システム構成ガイドを参照してください。

<http://www.nec.co.jp/products/express/systemguide/100guide.shtml>

3.6 注意事項

- ・ 本体装置内蔵物理デバイスに OS をインストールする時は、OS をインストールする物理デバイス以外は接続せず、OS インストール後に接続してください。

第4章 ソフトウェア編

4.1 RAID コントローラーのソフトウェア

RAID コントローラーを用いて RAID システムを構築するには、各 RAID コントローラーに対応したソフトウェアを使用する必要があります。RAID コントローラーを制御するソフトウェアは、オフラインユーティリティとオンラインユーティリティに大別されます。

4.2 オフラインユーティリティ

オフラインユーティリティは RAID コントローラーあるいは本体装置の BIOS ROM 内に格納されております。オペレーティングシステムを起動せずに RAID コントローラーの操作を行うことができます。

N コード/名称	オフラインユーティリティ名	略称
N8103-189/190/191/192/193/194/195/196/201/237/238	Smart Storage Administrator(オフライン版)	SSA
	システムユーティリティ	なし
オンボード RAID(SmartArray S100i)	Smart Storage Administrator(オフライン版)	SSA
	システムユーティリティ	なし

*各ユーティリティの操作方法については本体装置及び RAID コントローラーのユーザズガイドを参照してください。

4.3 オンラインユーティリティ

オンラインユーティリティは、オペレーティングシステムが起動した状態で、RAID システムの構築、RAID システムの監視を行うことができます。

N コード/名称	RAID システム管理ユーティリティ名	略称
N8103-189/190/191/192/193/194/195/196/201/237/238	Smart Storage Administrator	SSA
	RAID 通報サービス	なし
オンボード RAID(SmartArray S100i)	Smart Storage Administrator	SSA
	RAID 通報サービス	なし

4.4 RAID コントローラーのソフトウェアでの注意事項

[Smart Storage Administrator (オフライン版)]

- ・ POST 画面からの実行時にアプリケーションを終了する際は、本体装置の電源ボタンを長押しして電源を切断してください。
- ・ デバイスの構成およびステータスが更新されない場合は、画面上の[更新]ボタンをクリックするか、[F5]キーを押下してください。

[オンラインユーティリティ全般]

- ・ RAID コントローラーを使用する場合は、必ず RAID システム管理ユーティリティをシステムにインストールしてください。RAID システム管理ユーティリティをインストールしていない場合、RAID システムの障害検出ができません。また、RAID システム管理ユーティリティを使用する場合は、RAID システム管理ユーティリティの説明書を事前に読んでください。以下に示す注意事項以外にもシステムを運用するために留意すべき注意事項が記載されています。説明書は NEC サポートポータルサイト (URL <https://www.support.nec.co.jp/>)より入手できます。
- ・ RAID システム管理ユーティリティを使用する場合は、管理者権限のあるユーザでログオンしてください。管理者権限を持たないユーザでログオンした場合は、RAID システム管理ユーティリティが動作しない、または操作できない場合があります。

[Smart Storage Administrator]

- ・ デバイスの構成およびステータスが更新されない場合は、画面上の[更新]ボタンをクリックするか、[F5]キーを押下してください。
- ・ RAID コントローラーに接続しているドライブの場合でも、RAID を組まず OS より直接ドライブの設定や使用が可能です。ただし、それらドライブを使用して RAID を作成すると RAID で占有され管理下に置かれるため、該当ドライブおよびファイルが消失しますので、注意してください。

第5章 運用編

5.1 性能比較

5.1.1 性能比較

物理デバイスのタイプが性能に影響を与えます。

SSD はハードディスクよりも性能が高くなり、SAS デバイスは SATA デバイスよりも性能が高くなります。

また、一般的に RAID コントローラーはインテリジェントタイプがノンインテリジェントタイプよりも性能が高くなります。

5.2 RAID レベルの比

RAID コントローラーを用いて RAID システムを構築する前に、用途に応じた RAID レベルを選択する必要があります。耐障害性、アクセス速度および容量効率を考慮の上で最適な RAID レベルを選択してください。

①耐障害性

耐障害性は冗長構造を持つ RAID レベルを選択することで向上させることができます。RAID1 は構成に必要な物理デバイスの数が少なく、冗長構造も単純であるため耐障害性は高くなります。RAID5 は制御する物理デバイスの数が多くパリティ計算など、冗長構造が RAID1 よりも複雑であることから、耐障害性は RAID1 よりも低くなります。RAID6 は、2重のパリティを使用するため耐障害性は RAID5 より高くなります。

→「2.1.3 RAID の種類(RAID レベル)」

RAID1 > RAID6 > RAID5 RAID0 は耐障害性無し

[注意事項]

- 物理デバイスのマルチデッドによるシステム障害の発生を低減させる観点から、各アレイの物理デバイス搭載数は 8 台以下を目安とした RAID 構成を推奨します。
- 大容量物理デバイスにて RAID を構築する場合、障害復旧時に長時間のリビルドが必要です。その間冗長性が失われますので、より信頼性を高めるためにも物理デバイス 2 台の障害に対応する RAID6 あるいは RAID60 でのご利用を推奨します。

②アクセス速度

アクセス速度は物理デバイスを複数台接続しスプリットシークを行うことで向上させることができます。RAID0 はこの機能を用いたアクセス向上が期待される RAID レベルです。RAID5 と比べてパリティ情報の書き込みが必要ない分アクセス速度が速くなります。RAID1 と RAID5 との比較においてもライト動作においてパリティ計算+パリティライトが必要であるため RAID1 の方がアクセス速度は勝っています。RAID6 は、リード性能は RAID5 と同等ですが、ライト性能はパリティを2重化しているため低くなります。

→「2.1.3 RAID の種類(RAID レベル)」

RAID0 > RAID1 > RAID5 > RAID6

③容量効率

容量効率は物理デバイス総容量に占める冗長データ容量を少なくすることで向上させることができます。RAID1 と RAID5 は冗長構造を持ちます。RAID1 は2台の物理デバイスを用いて構築可能であり、小規模なシステム向けですが、容量効率は 50%に固定されます。RAID5 は3台以上の物理デバイスを用いて構築可能であり、中規模～大規模なシステム向けです。容量効率も 66%～と物理デバイスを多く接続すればするほど容量効率が上がる特徴を持っています。RAID6 は常に物理デバイス2台分の容量がパリティとして使用されるため、RAID5 と比べると容量効率は低くなり 33%～になります。

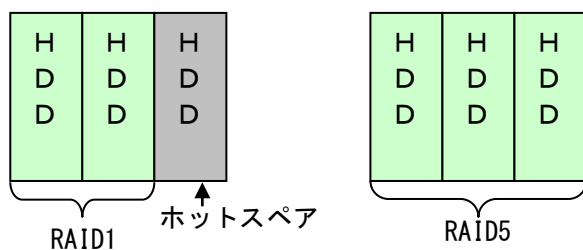
→「2.1.3 RAID の種類(RAID レベル)」

RAID0 > RAID5 > RAID6 > RAID1

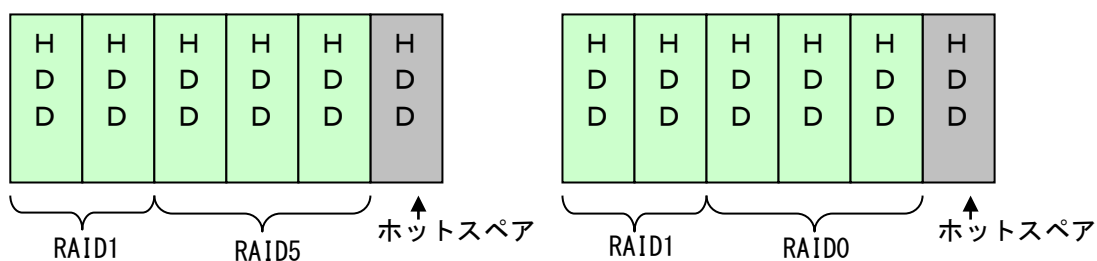
5.3 RAID システムの構築

上記、RAID コントローラーの選択および RAID レベルの選択をふまえ、RAID システムの構築例を下記に示します。

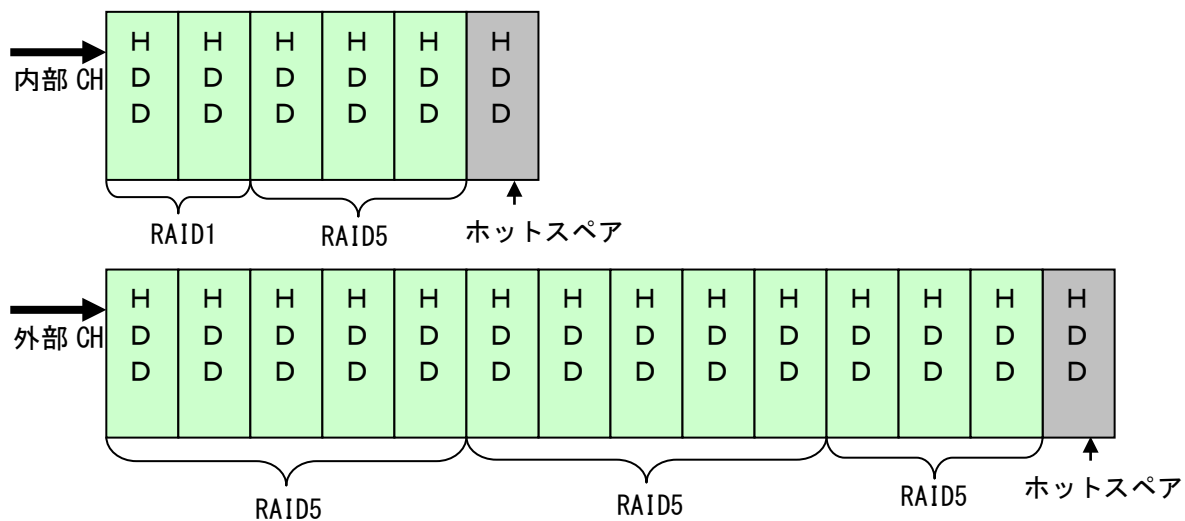
①小規模システム例（物理デバイス3台）



②中規模システム例（物理デバイス6台）



③大規模システム例（物理デバイス6台・内部ケーシング物理デバイス 14 台・外部 DISK 増設筐体）



5.4 安定運用のために

RAID コントローラーは RAID システムを構築し制御することで高速化、大容量化および高信頼性を提供します。ただし、RAID システムを構築するには複数の物理デバイスを必要とします。物理デバイスは技術の粋を結集した非常に高度な精密機械であり、デリケートな要素を持っています。RAID システムにおいて高信頼性を確保するには下記に示す操作を行う必要があります。

5.4.1 表面スキャンの実施

表面スキャンは、物理デバイスの後発不良に対する予防策として有効です。表面スキャンは、物理デバイスの使用領域をリードし、リードエラーを訂正します。これにより、物理デバイスの後発不良を予防することができます。

表面スキャン機能をサポートする RAID コントローラーを使用する場合は、表面スキャン機能を使用することを推奨します。

障害事例:

- (1)論理ドライブを構成する物理デバイスで故障(Dead)が発生。復旧のためにリビルドを行うと、故障(Dead)していない他の物理デバイスでリードエラーが発生し、リビルドが失敗。
- (2)論理ドライブを構成する物理デバイスで故障(Dead)が発生。縮退状態での運用でバックアップを実施すると故障(Dead)していない他の物理デバイスでリードエラーが発生し、バックアップが失敗。

原因:

複数の物理デバイスに後発不良または書き込み時の異常が発生していた場合、正常状態であれば冗長機能によりデータの復旧ができていたが、縮退状態ではデータを復旧できず、処理が中断する。

改善:

表面スキャンを有効にします。

- (1)表面スキャンの全面リードの際にエラー箇所を復旧するため、突然の物理デバイスの故障(Dead)が発生した場合でも安定稼働を継続することができます。
- (2)リビルド作業の失敗を低減します。表面スキャンの設定については、各 RAID コントローラーの RAID システム管理ユーティリティを参照してください。

5.4.2 RAID システム管理ユーティリティ+ESMPRO の利用によるアラート

障害事例:

- (1)1 台の物理デバイスの故障(Fail)に気が付かず、そのまま運用を続け、2ヶ月後に他の物理デバイスでエラーが発生し、2 台の物理デバイスが故障(Fail)してシステムダウン。

改善:

RAID システム管理ユーティリティ、および、ESMPRO を利用します。

- (1)RAID システム管理ユーティリティが故障(Fail)を監視しており、ESMPRO/ServerAgent は RAID システム管理ユーティリティが検知した物理デバイスの故障(Fail)情報を元に ESMPRO/ServerManager へアラートを送信します。システムダウンとなる重要障害を未然に防ぐことが可能です。(2 台の物理デバイスが故障(Fail)する前に交換/復旧可能)。
- (2)ESMPRO は RAID コントローラー配下の物理デバイス内部エラーのしきい値監視、S.M.A.R.T.監視を行い、物理デバイスが故障(Fail)する前にアラートを送信することが可能です。

注意:

ESMPRO/ServerManager,ESMPRO/ServerAgent による運用管理機能では、ESMPRO/ServerAgent が動作している被管理装置のアラートを、管理端末で動作する ESMPRO/ServerManager のアラートビューアで監視することができます。なお、オペレーションウィンドウ上ではステータスを確認することができません。

5.4.3 RAID システム管理ユーティリティの使用について

- ・ RAIDコントローラーを使用する場合は、必ずRAIDシステム管理ユーティリティをシステムにインストールしてください。RAIDシステム管理ユーティリティをインストールしていない場合、RAIDシステムの障害検出ができません。
- ・ RAIDシステム管理ユーティリティを使用する場合は、管理者権限のあるユーザ(administrator等)でログオンしてください。管理者権限を持たないユーザでログオンした場合は、RAIDシステム管理ユーティリティが動作しない、または操作できない場合があります。

5.4.4 RAID コントローラー用ドライバ、RAID システム管理ユーティリティのアップデート

既知の問題や機能強化を実施した RAID コントローラー用ドライバ、RAID システム管理ユーティリティの最新バージョンは、Web サイトに随時公開しています。定期的に確認し適時アップデートしてください。

NEC サポートポータルサイト (URL <https://www.support.nec.co.jp/>)

5.4.5 RAID 構成物理デバイス台数の設定による保守運用性の向上

障害事例:

1 台の RAID コントローラーに物理デバイス(73GB/1600rpm)を 16 台接続し、16 台の物理デバイス全容量を使用して 1 ディスクアレイ(RAID5)を構築した。整合性チェックを実施すると、約 18 時間を要し夜間作業にて整合性チェックが終了しなかった。そのため、定期的な整合性チェックが実施されることなく運用が継続され、障害発生時のリビルドにてリードエラーが起り、復旧に失敗した。

改善:

論理ドライブを構成する物理デバイスの台数を少なめに設定します。

(1)耐障害性・冗長性の改善

1 つのディスクアレイを構成する物理デバイス台数を少なくする(ディスクアレイを細分化させる)ことで、冗長性が向上(障害発生率が低下)します。

(2)保守運用性の改善

ディスクアレイを構成する物理デバイス台数を少なくすることで、リビルド時間も減少します。また、整合性チェックに要する時間も、RAID システムの構成を細分化することで、大幅に減少します。

[注意事項]

- ・ 物理デバイスのマルチデッドによるシステム障害の発生を低減させる観点から、各アレイの物理デバイス搭載数は8台以下を目安としたRAID構成を推奨します。
- ・ 大容量物理デバイスにてRAIDを構築する場合、障害復旧時に長時間のリビルドが必要です。その間冗長性が失われますので、より信頼性を高めるためにも物理デバイス2台の障害に対応するRAID6あるいはRAID60でのご利用を推奨します。