

仮想化環境に最適な アドバンストダイナミックプール



Advanced Dynamic Pool

1. はじめに

現状のビジネスを革新し、効率化するための手段として、クラウドコンピューティングが普及拡大しています。

「データの所有」から「データの使用」へと重点を移すこの大きな環境の変化の中で、さまざまな課題も出てきました。特にクラウド環境の基盤を構成するデータセンター内システムにおけるストレージについては、大きく分けて次の3つの課題が挙げられます。

1.1. データ量増加への対応

2011年の調査「IDC Digital Universe Study」によると、企業や個人が蓄積するデータは2年ごとに約2倍に増加し、2011年にはその総量は1.8ゼタバイト(=1.8兆ギガバイト)に達すると予測されました。^{*} しかも、こうした予測は毎年のように上方修正される傾向にあります。

このような環境の中で、ストレージ製品導入時に適切な容量設計を行うことは困難を極めます。

初期コストの削減を目的として小型ストレージ製品を導入し、データ量が予測を超えて増加した場合には、ストレージ容量の拡張や大型ストレージ製品への移行、またはデータの削除などの作業が発生します。

しかし、逆に余裕を待たせて当初から大型ストレージを導入した場合、今度は初期コストが増大します。また、データ量の増加が予測を下回った場合には、結果として無駄な設備投資を行ったことになります。

いかにして初期投資を抑制し、データ量増加に合わせて適切なストレージ容量の拡張を行っていくか。これが第一の課題です。

1.2. 管理・運用コストの抑制

従来、アクセス増加による性能低下や容量不足などの運用時に起こる問題に対して、個別最適化の方針で解決が図られてきました。しかし、その結果としてストレージ数が増加すれば管理業務の煩雑化を招き、ストレージが多様化すれば複雑化を招きます。また、消費電力などの運用コスト面から見て非効率的です。これらの要因により、管理・運用コストは年々増加する傾向にあります。

ストレージの空き容量が各ストレージに分散していることも問題です。利用状況に合わせて頻繁にデータ移行を行うことは現実的ではありません。大型ストレージ製品にデータを統合すればこうした問題は解決しますが、サーバ側を含めた設計作業が必要となるうえ、根本的な解決にはならないため、将来的に同じ問題が再発する可能性があります。

いかにして管理・運用コストを抑制し、柔軟にストレージの容量拡張・新規作成・削除を実現するか。これが第二の課題です。

1.3. システム停止時間の最小化

業務のグローバル化や企業内・企業間システムの連携により、システムの運用時間は長期化の一途をたどっています。ストレージは24時間365日、常に多数のサーバからアクセスされていると言ってもよいでしょう。

システム統合が進むと、部分的なシステムの停止がそのまま業務全体の遅延・停止につながりかねません。その結果として会社の社会的信用が低下する危険性すらありますから、停止する回数は可能な限り削減し、停止時間はできるかぎり短縮すべきです。

障害発生時やシステム構成の変更時にも、停止時間を最小化できる環境が求められています。

また、千年に一度と言われた2011年の大震災の教訓から、災害対策やデータのバックアップ体制にも注目が集まっています。

いかにしてシステムの運用を続け、利用者の無停止要求に応えるか。これが第三の課題です。

ストレージの課題

課題 1 データ量増加への対応

- 必要最小容量での導入 (スモールスタート)
- 柔軟なストレージ容量拡張

課題 2 管理・運用コストの抑制

- 管理の簡易化
- 運用コストの削減
- 余剰リソースの一元化

課題 3 システム停止時間の最小化

- 停止回数の最小化
- データのバックアップ体制

図 1：ストレージの課題

2. NEC が提案する解決策

そこで本稿では、根本的な解決策として、先進仮想化テクノロジーを採用した次世代ストレージ「iStorage Mシリーズ」の新機能「アドバンスダイナミックプール」について論じます。

ストレージ容量をあたかも液体のように扱い、システム運用中でも柔軟に論理ディスクの拡張・追加・削除を行うことを可能とした「ダイナミックプール」機能をさらに進化させた「アドバンスダイナミックプール」は、以下のように数多くの課題を解決します。

2.1. 課題 1「データ量増加への対応」への解答

● iStorage ThinProvisioningでスモールスタートを実現

iStorage ThinProvisioning(シンプロビジョニング)で容量を仮想化。実際に使用する領域だけを確保するため、導入時には必要最低限の実容量を搭載したスモールスタートが可能です。

● アドバンスダイナミックプールでHDDを追加するほど性能アップ

RAID方式に関わらず、HDDを1台ずつ何回でも増設可能。さらにアドバンスダイナミックプールの新機能「自律分散再配置」で、HDDの負荷を平準化し、性能を引き出します。

2.2. 課題 2「管理・運用コストの抑制」への解答

● ストレージはiStorageManagerで一元管理、さらに自動処理を設定してコストダウン

アドバンスダイナミックプールの機能は、すべてiStorageManagerでオンライン管理可能。自動処理に関するポリシーも細かく設定でき、運用・管理コストを削減できます。

● MAID機能で「データセンターまろごとエコ」

使用していないHDDはMAID機能で自動停止。さらに使用状況に応じて、ファンやCPUも自律制御します。消費電力を削減するとともに、発生熱量も減少するため、データセンターの空調費も削減できます。

● 空き容量はプールとして一括管理、論理ディスクに動的割り当て

各HDDの空き容量はプールとして集中管理。必要に応じて論理ディスクに容量を割り当てたり、動的に容量を拡張したりできるため、無駄なリソースを生みません。

2.3. 課題 3「システム停止時間の最小化」への解答

● 無停止でHDDを追加可能、しかも追加後はすぐに利用開始

HDDの追加・交換は、システムを停止せずに実施可能。しかもRAID構成に応じた台数を増設すれば、即時に利用を開始できます。

● RAID-1/TM/5/6に対応

アドバンスダイナミックプールでは、RAID-1/TM/5/6に対応。冗長性や容量効率を考慮して、最適なものを選択できます。

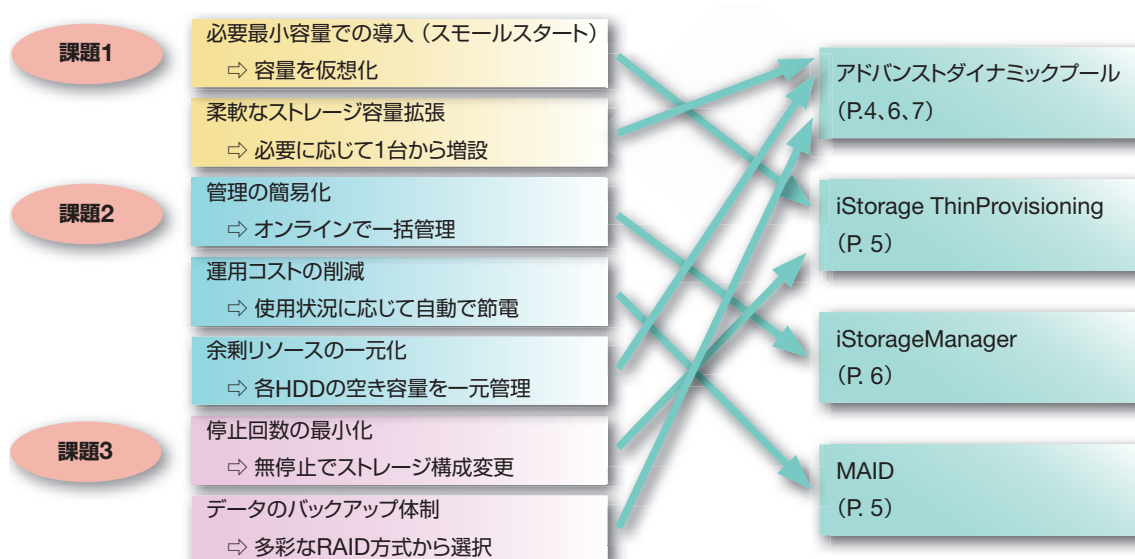


図 2 : NEC が提案する各課題への解決策

3. 課題を解決するさまざまな機能

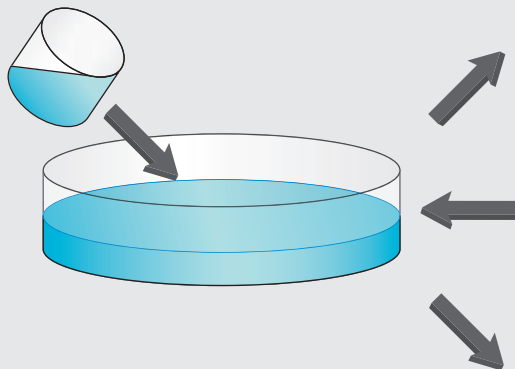
3.1. ダイナミックプール

ダイナミックプールの「プール」は文字通り水泳の「プール」を指します。ストレージ全体を大きなプールに見立て、水の量はHDDの合計容量を表します。水(=容量)が必要になったときにはプールから必要なだけ汲み出し(=論理ディスクの新規作成)、不要になったときにはふたたびプールに戻します(=論理ディスクの削除)。

プールの水が足りなくなれば、さらに水を増やす(=HDDの増設)こともできます。

HDD の増設

プールの容量が不足した場合、HDD を増設することでプールの容量を増やすことができます。



論理ディスクの新規作成

新たな論理ディスクが必要になった場合、プールから論理ディスクを新規作成できます。

論理ディスクの削除

論理ディスクが不要になった場合、プールに返却してプールの容量を増やすことができます。

論理ディスクの容量追加 *

論理ディスクの容量が不足した場合、プールから論理ディスクの容量を追加することができます。

*サーバから認識される容量を増加させる方法は、プラットフォームやファイルシステムによって異なります。

図3：ダイナミックプールの概念図

● HDDは1台から増設可能

容量が不足した場合、従来は例えばRAID 1であれば2台、RAID 5(4+P)であれば5台といったように、RAID方式に応じた台数のHDDを増設する必要がありました。

そこでダイナミックプールでは、HDDの物理構成とRAID制御を切り離し、RAIDを仮想化しました。HDDを増設することに台数に応じてデータを再構築する「分散マッピング方式」の採用により、RAID方式に関係なく、必要に応じてHDDを1台ずつ何度でも増設できます。

つまり、ストレージの物理構成に縛られることなく、柔軟な運用を実現できます。

3.2. iStorage ThinProvisioning

「シンプロビジョニング」はストレージの容量を仮想化する技術です。iStorage ThinProvisioningは、サーバに実容量よりも大きな仮想容量を認識させ、実際には書き込まれたデータに応じて必要な容量のみを動的に領域を確保します。この仕組みにより、通常は導入時に必要となる容量設計が不要になります。

例えば、導入時に将来のデータ量増加を見越して大きめの仮想容量10TBを業務サーバに認識させ、実際は使用する2TBのみを割り当てておくといったスモールスタートが可能です。その後、データの増大に応じて、運用しながらストレージを拡張します。

また、しきい値を設定して自動で動的に容量拡張を行わせることも可能です。

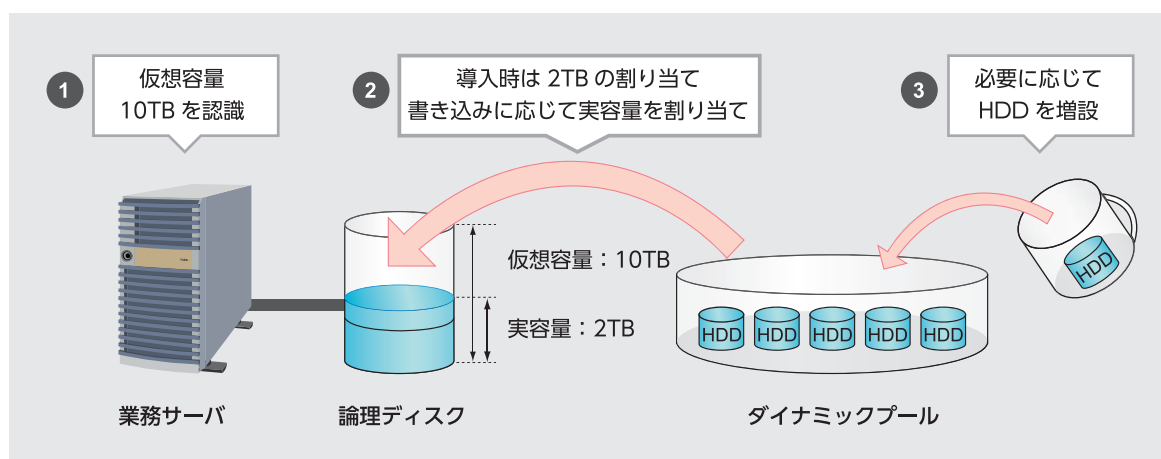


図4：シンプロビジョニングの概念図

また、シンプロビジョニングはデータのバックアップ／リストア時にも利用されます。通常、バックアップは全領域をコピーするため、実際にはデータが存在しない領域まで容量が確保されてしまいます。iStorage ThinProvisioningはゼロデータ（＝データがオール0）の書き込みを検出する仕組みもっており、最小の容量でデータのバックアップ／リストアを行うことができます。

3.3. MAID による消費電力削減

NECは、2007年度からITプラットフォームの省電力化に取り組んでおり、特にサーバやストレージの消費電力の削減に徹底して取り組んできました。

MAID (Massive Arrays of Inactive Disks) もそうした取り組みのひとつで、ストレージ内で未使用状態のHDDの回転を停止し、消費電力を削減する機能です。iStorageではDシリーズからこの機能を実用化しました。

プール中で未使用状況になっているHDDの回転を自動的に停止し、データ量の増加などで必要となった場合に自動的に再稼働させることができます。

また、スベア用のHDDの回転を停止しておき、障害発生時に自動的に再稼働するように設定することもできます。停止設定はiStorageManagerで行うため、追加ソフトは必要ありません。

3.4. 消費電力の「見える化」で MAID 機能をさらに強化

同一インフラ内のすべてのiStorage Mシリーズの消費電力は、リアルタイムでビジュアル表示されます。

また、ストレージの利用状況に合わせてCPU、ファンなどを自律制御し、消費電力を削減します。

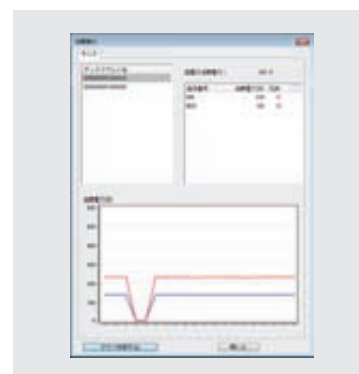


図5：消費電力の見える化

3.5. iStorageManager による一元管理

iStorageManagerは、論理ディスクや、ディスクアレイを構成するHDD、コントローラ、電源などの状態を一元管理するソフトウェアです。Webブラウザから直感的な操作で運用できます。

障害が発生した場合や、設定した各種しきい値を超えた場合には、メッセージやアイコンで通知します。

iStorageManager Expressは、サーバ部がディスクアレイにインストールされているため、管理サーバなしで1台のディスクアレイを効果的に管理できます。

また、WebSAM iStorageManagerを使用すれば、複数台のディスクアレイを一括管理可能です。



図 6 : iStorageManager の画面イメージ

3.6. すべての作業はオンラインで実行可能

こうしたアドバンスダイナミックプールの機能や操作は、すべてオンラインで実行できます。またHDDの増設後の処理などはバックグラウンドで行われ、日中は業務に与える影響を10%以下にとどめ、夜間など業務負荷がない時間帯には処理を加速するように設定されています。

さらに業務処理とバックグラウンド処理のバランスは自由に設定できるため、運用状況に応じたバックグラウンド処理が可能です。

4. アドバンスダイナミックプールの新機能

iStorage Mシリーズに搭載されるアドバンスダイナミックプールではダイナミックプールの概念をさらに進化させ、増設時に自動実行されるデータの再配置アルゴリズムを改良し、パフォーマンスの向上を果たしました。

4.1. 自律分散再配置によりパフォーマンスを向上

HDD増設後はデータを均等に配置して負荷を均等化する「自律分散再配置」を実施し、パフォーマンスを向上させます。

図7はRAID 1を構成するHDDが2組(=4台)存在した場合の図です。アドバンスダイナミックプールでは、いったんデータを読み出し、メモリに展開した後、右図のようにRAID 10/4台構成へと再配置します。

同時に、各HDDのデータ量を平準化し、運用時の負荷の均等化を図ります。

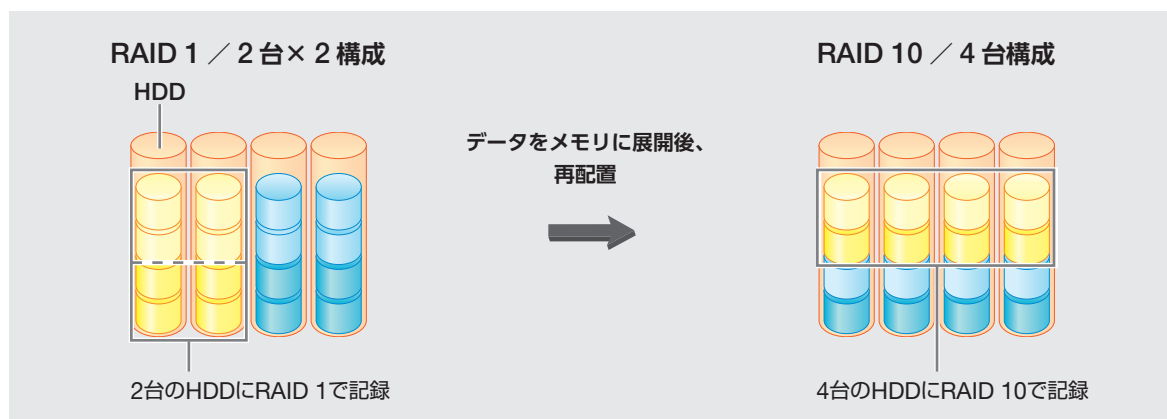


図 7 : 自律分散配置の概念図

4.2. HDD の数に応じて自動でストライプを作成

アドバンスダイナミックプールでは、RAID (Redundant Array of Independent Disks)構成を採用してHDD の故障によるデータ消失を防ぐことができます。RAID 1、TM、5、6 に対応しており、冗長性や容量効率などの特性を考慮して、お客様の使い方に最適な方式を選択可能です。

さらにHDDの台数に応じてバックグラウンドでデータの再配置を行い、自動的にストライプが作成されます。

RAID 方式	ストライプ作成条件
RAID 1	4台以上でRAID 10
RAID TM(トリプルミラー)	6台以上でRAID TM × n
RAID 5(4+P)	10台以上でRAID 50/(4+P) × n
RAID 6(4+PQ)	12台以上でRAID 60/(4+PQ) × n

図 8 : RAID 方式とストライプ作成条件

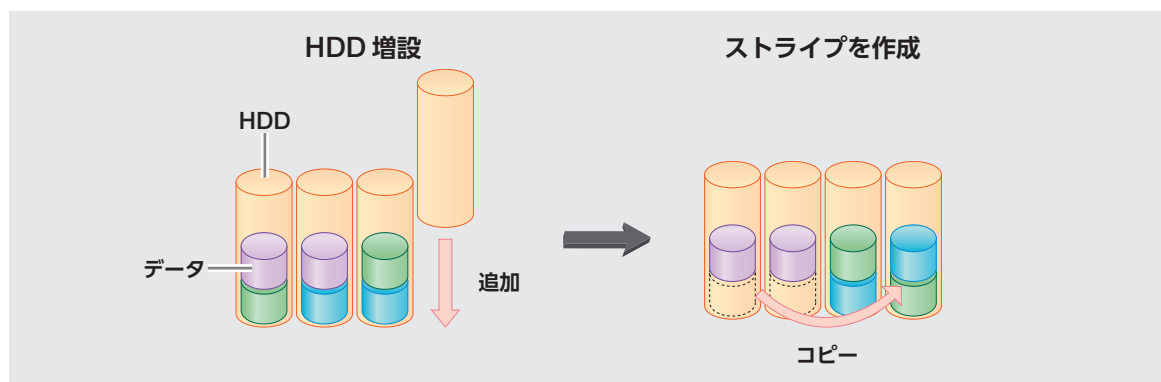


図 9 : ストライプ作成の概念図 (RAID 1、HDD を 3 台から 4 台に増設したケース)

4.3. データサイズに応じた配置

ストライプを作成する際は、複数のRANK(RAIDグループ)にまたがってデータを配置することで、複数のHDDに並列にアクセスできるようにし、パフォーマンスを向上させます。

しかし、サイズの小さなデータを多数のHDDに分散させると、それぞれのHDDにアクセスする際に発生するシークタイムや回転待ちなどの影響により、パフォーマンスが逆に低下します。そこで、1MB以下のデータはRANK内で完結するように配置します。

従来機ではRAID 10時にもRAID 1 相当のパフォーマンスにとどまる場合がありますでしたが、この配置技術により性能を確保しています。

また、HDDが故障したときにはRANKを構成する他のHDDすべてを読み出す必要があるため、小さいデータを分散させないことで故障時の性能低下を最小化できるという効果もあります。

5. 実証例

5.1. 自律分散配置の効果

第4.1項で触れた自律分散配置による性能向上を図10に示します。HDDの増加に従って直線的に性能が向上していることがご覧になれます。

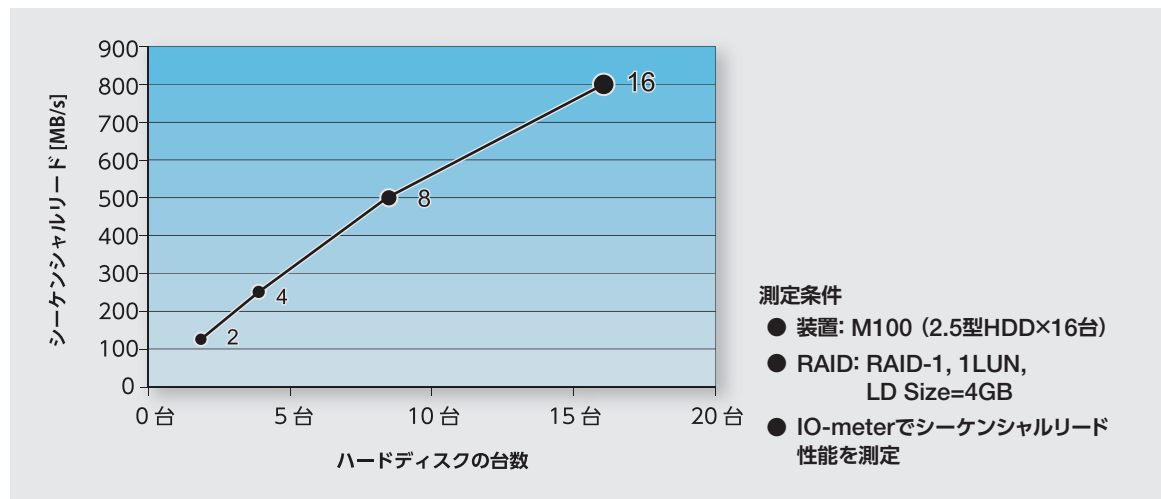


図 10 : HDD 増設による性能向上

5.2. MAID 機能による消費電力削減効果

図11は運用中のHDD24台のうち、20台を休止させたケースです。70W以上の電力の削減を達成しました。



図 11 : MAID 機能による消費電力削減効果

※出典: IDC White Paper, Extracting Value from Chaos, June 2011, Sponsored by EMC Corporation

お問い合わせは、下記へ

NEC プラットフォーム販売本部(ストレージお問い合わせ)

〒108-8424 東京都港区芝五丁目33番8号(第一田町ビル)
TEL: 03(3798)9740

問い合わせURL <http://www.nec.co.jp/contact>
国内向け製品URL <http://www.istorage.jp>

人と地球にやさしい情報社会へ

●本ホワイトペーパー内の社名、商品名は各社の登録商標または商標です。
●製品の仕様は、都合により予告なしに変更することがあります。