



SmartSyncバックアップ – 効率的NAS-to-NASバックアップ

1. 要約

NASデータをバックアップする共通のアプローチは、Windows やUNIX端末上でバックアップソフトウェアを走らせ、ネットワークを介してNASをバックアップするというものです。バックアップ時間は、ネットワークトラフィック、NASのパフォーマンス、テープドライブ速度に依存します。大量のテープカセット、インクリメンタル(増・差分)バックアップ、フルバックアップ、複数のテープに跨るバックアップなどを扱う際には、データのリストアもいささか複雑になります。

SmartSyncバックアップは、NAS Storageサーバ上で実行します。ファイル全体ではなく、変更されたデータブロックのみ転送することにより、バックアップスピードの改善を図っています。一方、常にオンライン上にフルバックアップを構築し、オンライン上にそれを保持することにより、データのリストアを容易にしています。

2. SmartSync バックアップの設計コンセプト

現在ほとんどのNASバックアップは、バックアップソフトウェアを走らせたバックアップサーバに頼っています。このバックアップソリューションは一般的ですが、まだいくつかの面で完璧ではありません。SmartSyncバックアップソフトウェアは、同一同期の手法との組み合わせにより、NASバックアップを改良し、従来のバックアップソリューションを補完しています。SmartSyncバックアップソフトウェアが取り扱おうとしている課題は、以下の通りです。

● バックアップ時間の長さ

バックアップ時間は多くの要因に影響を受けます - NASからソースデータをどれくらい速く読み出すことができるか、どれくらいのネットワークトラフィックがあるか、1分間に何メガバイトをテープに書き込むことができるか、など。IT技術者は、いつもバックアップスピードを改善し、バックアップ時間を短くし、営業活動への影響を最小限にしようとします。選択肢があまり無いため、ハイエンドのハードウェア、ハイスピードのネットワーク、高速なテープ装置に投資し続けています。

ハイエンドのハードウェア / ソフトウェアの購入やネットワークインフラの更新に大金を支払うこと他に、何かできることはないのでしょうか？

SmartSyncバックアップソフトウェアは、バックアップスピードを根本から改善しようとしています - そのアルゴリズムから。

従来のバックアップソフトウェアは、数バイトが変更されただけでもファイル全体をバックアップ作成時、テープにコピーします。ほとんどの場合、ファイルが大幅な変更を受けることはあまり無いため、大量のリソースが無駄に消費されてしまいます。異なるバージョンのファイルは、たとえ頻繁に変更されたとしても、約70%の共通部分を持っています。このファイルのバックアップを毎日行った場合、70%の同一部分を含む全体ファイルを毎日転送することになります。どうして30%の差異部分だけを転送しないのでしょうか？

NS-ELEX

SmartSyncバックアップの発想は、最終バックアップからの変更部分のみを転送するというものです。ファイルの一部だけ転送されればよいので、バックアップスピードは大幅に改善され、さらに、バックアップ中のネットワークトラフィックへの要求は減りますので、ネットワークを利用する業務への影響は最小で済みます。

• データリストアの困難さ

一般的に、データがクラッシュした際は、IT技術者はデータを復旧するために、最新のフルバックアップといくつかのインクリメンタルバックアップデータをリストアしなくてはなりません。その処理過程は複雑です。まず、データのある場所を探し出し、それらのファイルをどのテープが保持しているかを探し出す必要があります。次に、オフラインに収納されていたり、同じサイトに収容されていたり、なかったりするテープを見付けなくてはなりません。そして、2回以上データリストアを実行します

- 1回目はフルバックアップのため、2回目以降はインクリメンタルバックアップのためにです。

SmartSyncバックアップは、すべてのバックアップをオンラインに保持します。さらに良いことに、常にフルバックアップの状態を作り出し、保持しています。IT技術者は、データリストアを1回実行するだけです - 1つのバックアップバージョンをピックアップし、リストアするだけです。すべてのタスクはオンラインで行うことができ、分かりやすく簡単です。

• セキュリティ上の問題

バックアップデータはネットワークを介してバックアップサーバへ送られるため、セキュリティリスクが伴います。SmartSyncバックアップは、強力な暗号でデータを包み込み、盗聴を防止しています。誰かがデータパケットを傍受したとしても、その内容を理解することは不可能です。

高速バックアップ。フルバックアップ。オンラインバックアップ。これらがSmartSyncバックアップソフトウェアの設計コンセプトです。その目的は、バックアップ時間を縮めることや、データリストアを容易にすることです。これを可能にしている技術がSmartSyncアルゴリズムであり、差分ブロック転送を利用し、ネットワーク帯域幅への要求を減じてバックアップ時間を大幅に短縮しました。今やいつでもフルバックアップの状態を存在させることが実現可能になりました。MIS (経営情報システム (Management Information System)) 技術者は、テープのピックアップ、ファイルの検索、複数バックアップのリストアといった複雑な作業に対処する必要はありません。ただ1つのバックアップを選択し、リストアするだけです。データリストアは簡略化されました。

3. SmartSyncバックアップの動作

SmartSyncバックアップは、NAStorageからNAStorageへのバックアップを行います。

1台がSmartSyncバックアップサーバとして、もう1台もしくは複数台がSmartSyncバックアップクライアントとして動作しますので、2台以上のNAStorageシステムが必要になります。SmartSyncバックアップサーバが、クライアントデータのバックアップを構築し、保持します。

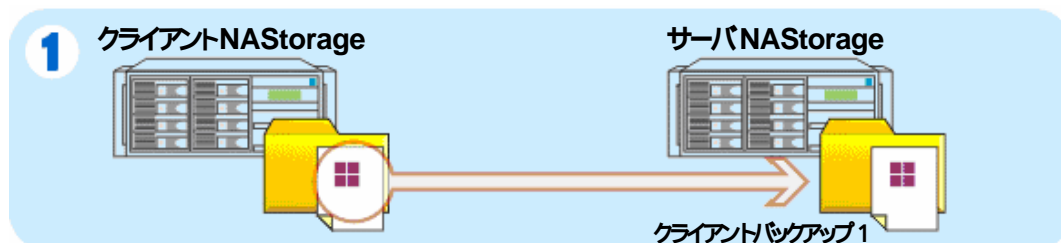
(注：すべてのNAStorageモデルはSmartSyncバックアップクライアントになり、一方、NAStorage 8200以上のモデルがSmartSyncバックアップサーバになります。)

ファイルは、データブロックから成ります。データをバックアップするには、クライアントはSmartSyncバックアップサーバにデータブロックを送ります。SmartSyncバックアップの内部動作とその高効率

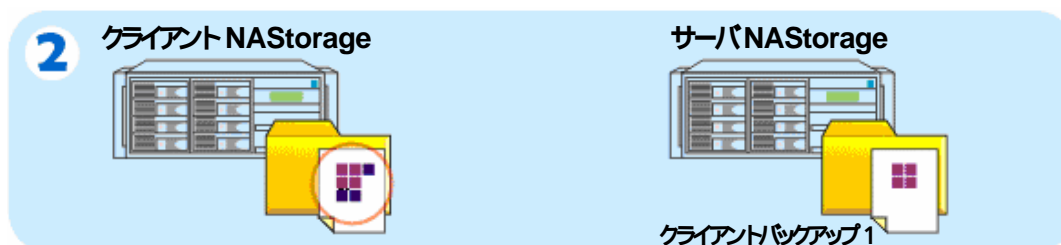
NS-ELEX

性を、イラストを使って紹介しましょう。

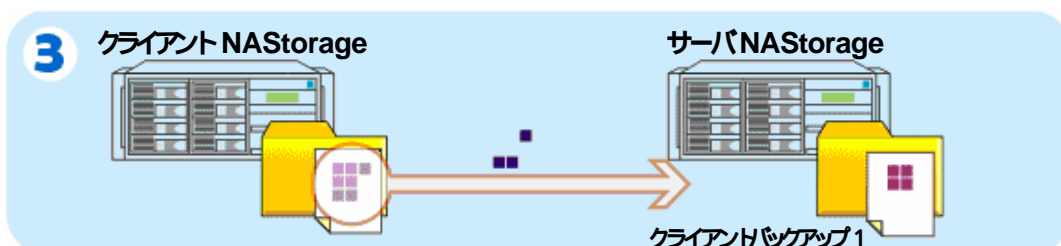
1. 最初は、クライアントデータのバックアップを、SmartSyncバックアップサーバに送ります。



2. クライアントデータに、いくつかの変更が加わります。



3. SmartSync/バックアップタスクが始まります。SmartSync/バックアップサーバに、変更されたブロックを転送します。



4. SmartSync/バックアップサーバはデータを再構成し、受け取ったブロックとクライアントバックアップ1に基づき、新規バックアップであるクライアントバックアップ2を生成します。



新規バックアップが、いくつかのデータブロックのみの転送で生成されたことが分かります。

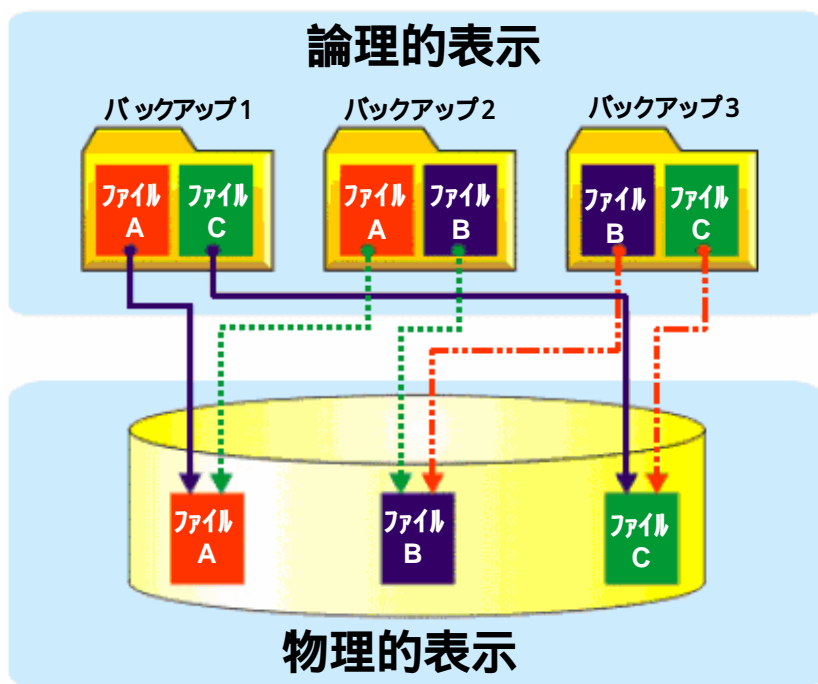
NS-ELEX

4. ディスク容量を節約するアイデア

- ファイルのワン・コピーのみ保持

SmartSyncバックアップは、常にフルバックアップの状態を作り出し、それをオンライン上に保持しています。大量のディスク容量が必要だと批評する人がいるかもしれませんが、バックアップからバックアップへ変更されないままの多くのファイルがあることは、珍しいことではありません。異なるバックアップバージョンに同じファイルのコピーを保持すると、ディスク容量を無駄に消費します。

SmartSyncバックアップソフトウェアはハードリンクを利用することにより、この問題を解決しました。ハードディスク上には、各ファイルのワン・コピーのみを保持します。ファイルシステムレベルでは、異なるバックアップバージョンは、ファイルを指し示すハードリンクを使用しています。下図を参照してください。



- ディスク容量を自動的にクリーンナップ

SmartSyncバックアップは、バックアップサーバにすべてのバックアップを保持しています。適切な管理を行わないと、バックアップサーバは思いがけなく容量不足を引き起こすでしょう。古くなったバックアップを手動で削除することもさることながら、SmartSyncバックアップは、事前に定義した方針に基づき、バックアップを削除するメカニズムを搭載しています。このバックアップバージョン制御は、テープバックアップで広く一般に使用されているGFS (Godfather-Father-Son) テープローテーションスキームにより着想されたものです。

NS-ELEX

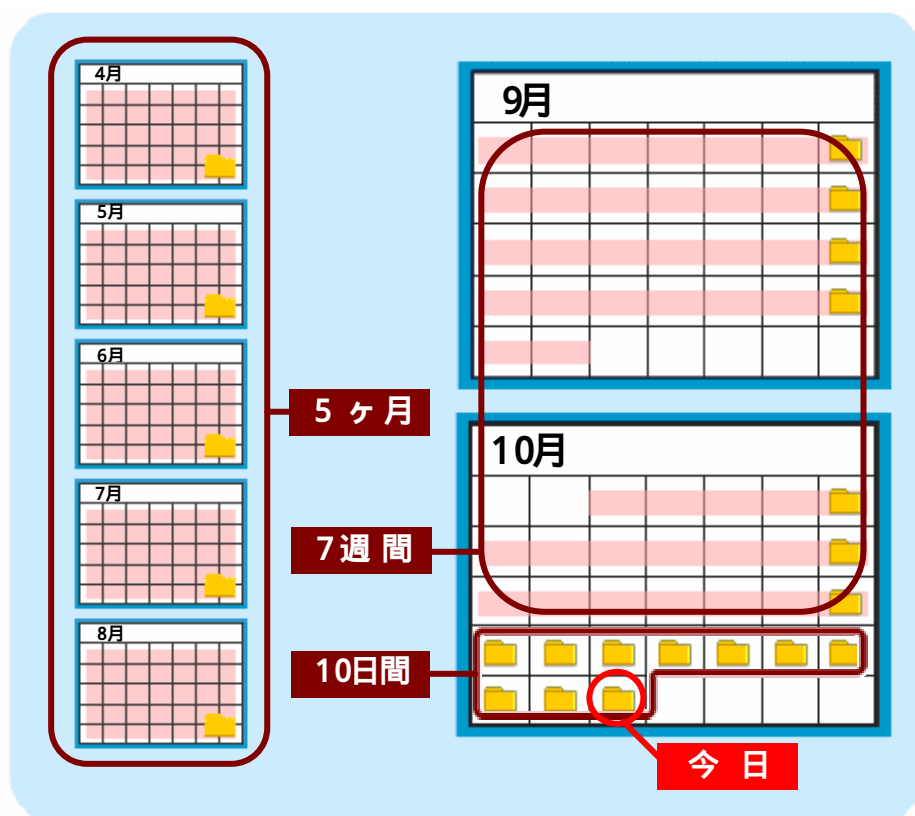
これを **拡張 GFSメディアローテーションスキーム**と呼んでいて、テープローテーションにおいて見られるようなテープ再利用の代わりに、ここでは将来利用のためにハードディスクを使えるようにします。新しいバックアップバージョンが作成されると、ルールに従い古くなったバックアップを自動的に照査し、削除します。

ルールは以下に述べるとおりです。X、Y、Zは、ユーザが定義した数字です。

- (1) 最近のX日間、毎日 1つのバックアップバージョンを保持 - 1日に2つ以上のバックアップバージョンがある場合、一番新しいバックアップのみ残り、他は削除されます。
- (2) X日間より前のY週間、1週につき1つのバックアップを保持 - 各週で一番新しいバックアップを保持します。
- (3) Y週間より前のZヶ月間、1月につき1つのバックアップを保持 - 各月で一番新しいバックアップを保持します。

ここで、以下のケースを図解してみましょう。

X=10 (10日間)、Y=7 (7週間)、Z=5 (5ヶ月間)とした場合、バックアップサーバに残るバックアップバージョンは以下のようになります。



NS-ELEX

5. SmartSync バックアップ対テープバックアップ

• バックアップメディアの単価

テープバックアップは、以前はGBあたりの価格が安いと公言されてきました。しかしながら、ハードディスクの価格が安価になり、より魅力的になってきました。

以下の価格は2003年10月に調査した参考平均価格です。

| | SuperDLTテープ 110GB 非圧縮容量 | LTOテープ 100GB 非圧縮容量 | Maxtor DiamondMax Plus 9 200GB (7200rpm) |
|----------|----------------------------|-----------------------|--|
| 価格 | 約 ¥11,000- | 約 ¥5,500- | 約 ¥24,000- |
| GBあたりの価格 | ¥1,000- | ¥550- | ¥240- |

HDDの単価は、SuperDLTテープとほとんど同じであることが分かります。メンテナンスコストと管理コストを考えると、テープは最低価格のメディアとは限りません。

• バックアップ時間

上述したように、SmartSyncバックアップソフトウェアは、バックアップ中にファイルの変更されたブロックのみを送ります。ファイルに変更があるのはたいてい一部分であるため、バックアップ時間は大幅に削減されます。

初回バックアップの場合には、SmartSyncバックアップソフトウェアは、基準バックアップのためにすべてのデータを送らなくてはなりません。実測テストでは、SmartSyncは14MB/s (840MB/分) でデータを転送できました。参考までに、某社SuperDLTテープドライブは11MB/s (660MB/分) でデータを転送しました。

• オンライン、オフライン、オフサイトデータバックアップ

SmartSyncバックアップは、すべてのバックアップをオンラインに保持します。一方、テープは常にオフラインです。テープ上にあるファイルを直接開くことはできません。ファイルを読むためには、バックアップテープをリストアしなければなりません。

しかしながら、テープを使わなければならない状況も存在します。恒久的データロスやシステムダメージを引き起こす人為的ミスや自然災害を回避するため、オフサイトにバックアップメディアを保存しておきたい場合には、テープバックアップは最良の選択です。

このようにデータをオフラインで保存する場合には、SmartSyncバックアップは役に立ちません。

• 複数クライアントの同時バックアップ

テープメディアはシーケンシャルアクセスしかできません。同時に1つのテープメディアに複数のバックアップを書き込むことは不可能です。SmartSyncバックアップはバックアップメディアとしてハードディスクを使用するので、同時アクセスが可能です。複数のSmartSyncバックアップタスクを同時



に実行できます。実際、SmartSyncバックアップサーバでは、1つのSyncポイントに対して最大 8 つの同時バックアップが可能です。

6. SmartSyncバックアップの利点

SmartSyncバックアップは、従来のテープバックアップとは多くの点で異なります。テープバックアップを置き換えるものではなく、テープバックアップを補完するものです。オンラインのバックアップデータを好むのであれば、SmartSyncバックアップは良い選択肢でしょう

SmartSyncバックアップの利点：

- ・ インクリメンタルバックアップより短いバックアップ時間でフルバックアップ可能
- ・ 差分ブロック転送により高速バックアップが可能
- ・ すべてのバックアップはオンラインにあり、リストアが簡単
- ・ すべての通信とデータ転送は暗号化され、安全
- ・ バックアップバージョン管理が可能