

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2008-529187

(P2008-529187A)

(43) 公表日 平成20年7月31日(2008.7.31)

(51) Int.Cl.

G06F 3/06 (2006.01)

F I

G06F 3/06 304F

テーマコード (参考)

5B065

審査請求 未請求 予備審査請求 未請求 (全 17 頁)

(21) 出願番号 特願2007-554068 (P2007-554068)  
 (86) (22) 出願日 平成17年8月31日 (2005. 8. 31)  
 (85) 翻訳文提出日 平成19年9月18日 (2007. 9. 18)  
 (86) 国際出願番号 PCT/US2005/030847  
 (87) 国際公開番号 W02006/083327  
 (87) 国際公開日 平成18年8月10日 (2006. 8. 10)  
 (31) 優先権主張番号 11/050, 239  
 (32) 優先日 平成17年2月3日 (2005. 2. 3)  
 (33) 優先権主張国 米国 (US)

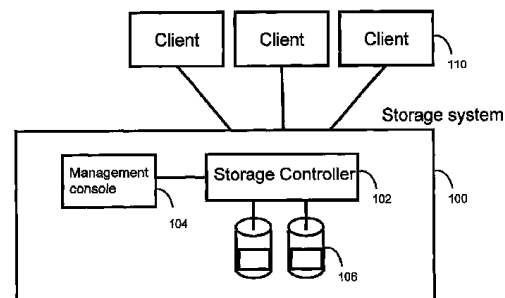
(71) 出願人 390009531  
 インターナショナル・ビジネス・マシーンズ・コーポレーション  
 INTERNATIONAL BUSINESS MACHINES CORPORATION  
 アメリカ合衆国10504 ニューヨーク州 アーモンク ニュー オーチャードロード  
 (74) 代理人 100108501  
 弁理士 上野 剛史  
 (74) 代理人 100112690  
 弁理士 太佐 種一  
 (74) 代理人 100091568  
 弁理士 市位 嘉宏

最終頁に続く

(54) 【発明の名称】 新規のポイント・イン・タイム・コピー動作

## (57) 【要約】

方法およびサービスは、ソース・ストレージ・ユニット内に記憶されたソース・データの仮想ポイント・イン・タイム・コピーを作成および保持する。当該方法/サービスは、ソース・データのポイント・イン・タイム・コピーを作成するための少なくとも1つの要求を受信する。しかしながら、ソース・データのコピーを作成する代わりに、本発明は、対象ストレージ・ユニットまたは同一のストレージ・システムによって管理される他のストレージ・ユニット内の対象ユニット・マッピング・テーブルを作成する。このターゲット・ストレージ・ユニット・マッピング・テーブルは、ソース・データに対するポインタを含む。加えて、本発明は、対象ストレージ・ユニットまたは同一のストレージ・システムによって管理される他のストレージ・ユニット内の修正空間を保持する。修正空間の各部分は、所定の対象ストレージ・ユニットと関連付けられている。修正空間は、対応する対象ストレージ・ユニットに固有のソース・データに対する変更を記憶するのみである。対象ストレージ・ユニット・マッピング・テーブルは、修正空間にデータが書



**【特許請求の範囲】****【請求項 1】**

ソース・データの仮想ポイント・イン・タイム・コピーを保持するためのコンピュータ・システムであって、

前記ソース・データに対するポインタを保持するように適合されたソース・ストレージ・ユニット・マッピング・テーブルと、

前記ソース・データに対するポインタと、修正データに対するポインタとを保持するように適合された少なくとも 1 つの対象ストレージ・ユニット・マッピング・テーブルとを備え、

前記対象ストレージ・ユニット・マッピング・テーブル内のポインタが対応するソース・データの代わりに前記修正データをポイントする以外は、前記対象ストレージ・ユニット・マッピング・テーブルは、前記ソース・ストレージ・ユニット・マッピング・テーブルと同一である、システム。

10

**【請求項 2】**

ある期間に渡って、前記対象ストレージ・ユニット・マッピング・テーブルは、前記ソース・データの部分と、前記修正データの部分とをポイントするポインタの固有の組み合わせを通じて、前記ソース・データの固有の仮想コピーを保持する、請求項 1 に記載のシステム。

**【請求項 3】**

前記ポインタは、直接ポインタおよび間接ポインタのうちのいずれか 1 つを備える、請求項 1 に記載のシステム。

20

**【請求項 4】**

前記修正データを保持するように適合される修正空間をさらに備える、請求項 1 に記載のシステム。

**【請求項 5】**

前記修正空間として利用可能なストレージの空きブロックのリストをさらに備える、請求項 4 に記載のシステム。

**【請求項 6】**

前記ソース・ストレージ・ユニット・マッピング・テーブルは、前記ソース・データを保持するソース・ストレージ・ユニット内の物理ストレージ装置上の位置に対するポインタを含む、請求項 1 に記載のシステム。

30

**【請求項 7】**

前記ソース・ストレージ・ユニット・マッピング・テーブルおよび前記対象ストレージ・ユニット・マッピング・テーブルは、それぞれ、

ポインタのマッピング・テーブル、

フラグのマッピング・テーブル、および

ハッシング・テーブルを伴うリンク付けされたリスト

のうちの 1 つを備える、請求項 1 に記載のシステム。

**【請求項 8】**

ソース・データの仮想ポイント・イン・タイム・コピーを保持するためのコンピュータ・システムであって、

40

少なくとも 1 つの物理ストレージ装置を備えるソース・ストレージ・ユニットと、

前記ソース・ストレージ・ユニット内の前記ソース・データに対するポインタを保持するように適合されたソース・ストレージ・ユニット・マッピング・テーブルと、

対象ストレージ・ユニット・マッピング・テーブルを備える少なくとも 1 つの対象ストレージ・ユニットとを備え、前記対象ストレージ・ユニット・マッピング・テーブルは、前記ソース・ストレージ・ユニット内の前記ソースデータに対するポインタと、修正空間内に保持された修正データに対するポインタとを保持するように適合される、システム。

**【請求項 9】**

ある期間に渡って、前記対象ストレージ・ユニット・マッピング・テーブルは、前記ソ

50

ース・データの部分と、前記修正データの部分とをポイントするポインタの固有の組み合わせを通じて、前記ソース・データの固有の仮想コピーを保持する、請求項 8 に記載のシステム。

【請求項 10】

前記ポインタは、直接ポインタおよび間接ポインタのうちのいずれか 1 つを備える、請求項 8 に記載のシステム。

【請求項 11】

前記修正空間として利用可能なストレージの空きブロックのリストをさらに備える、請求項 8 に記載のシステム。

【請求項 12】

前記ソース・ストレージ・ユニット・マッピング・テーブルは、前記ソース・ストレージ・ユニット内の前記物理ストレージ装置上の位置に対するポインタを含む、請求項 8 に記載のシステム。

【請求項 13】

前記ソース・ストレージ・ユニット・マッピング・テーブルおよび前記対象ストレージ・ユニット・マッピング・テーブルは、それぞれ、ポインタのマッピング・テーブル、フラグのマッピング・テーブル、およびハッシング・テーブルを伴うリンク付けされたリストのうちの 1 つを備える、請求項 8 に記載のシステム。

【請求項 14】

ソース・ストレージ・ユニット内に記憶されたソース・データの仮想ポイント・イン・タイム・コピーを作成および保持する方法であって、

前記ソース・データのポイント・イン・タイム・コピーを作成するための要求を受信するステップと、

前記ソース・データの前記コピーを作成する代わりに、前記ソース・データに対するポインタを含む対象ストレージ・ユニット・マッピング・テーブルを作成するステップと、

前記対象ストレージ・ユニットに関連付けられている修正空間であって、前記対象ストレージ・ユニットに固有の前記ソース・データに対する変更を記憶するだけである前記修正空間を保持するステップと、

前記修正空間に対してデータが書き込まれながら、前記対象ストレージ・ユニット・マッピング・テーブルを修正するステップであって、前記対象ストレージ・ユニット・マッピング・テーブル内の対応するポインタを前記ソース・データから前記修正空間へ転送するステップと

を含む、方法。

【請求項 15】

ある期間に渡って、前記対象ストレージ・ユニット・マッピング・テーブルは、前記ソース・データの部分と、前記修正空間の部分とをポイントするポインタの固有の組み合わせを通じて、前記対象ストレージ・ユニットのための前記ソース・データの固有のコピーを保持する、請求項 14 に記載の方法。

【請求項 16】

前記ポインタは、直接ポインタおよび間接ポインタのうちのいずれか 1 つを備える、請求項 14 に記載の方法。

【請求項 17】

前記修正空間を設けるステップをさらに含む、請求項 14 に記載の方法。

【請求項 18】

前記ソース・ストレージ・ユニット・マッピング・テーブルは、前記ソース・ストレージ・ユニット内の物理ストレージ装置上の位置に対するポインタを含む、請求項 14 に記載の方法。

【請求項 19】

10

20

30

40

50

前記ソース・ストレージ・ユニット・マッピング・テーブルおよび前記対象ストレージ・ユニット・マッピング・テーブルは、それぞれ、ポインタのマッピング・テーブル、フラグのマッピング・テーブル、およびハッシング・テーブルを伴うリンク付けされたリストのうちの1つを備える、請求項14に記載の方法。

【請求項20】

ソース・ストレージ・ユニット内に記憶されたソース・データの仮想ポイント・イン・タイム・コピーを作成および保持するための方法であって、

前記ソース・データのポイント・イン・タイム・コピーを作成するための少なくとも1つの要求を受信するステップと、

前記ソース・ストレージ・ユニットをソース・ボリュームとしてマーク付けするステップであって、前記ソース・ストレージ・ユニットがソース・ボリュームとしてマーク付けされる場合には修正はできない、ステップと、

前記ソース・データの前記コピーを作成する代わりに、前記ソース・データに対するポインタを含む対象ストレージ・ユニット・マッピング・テーブルを作成するステップと、

前記対象ストレージ・ユニットに関連付けられている修正空間であって、前記対象ストレージ・ユニットに固有の前記ソース・データに対する変更を記憶するだけである前記修正空間を保持するステップと、

前記修正空間に対してデータが書き込まれながら、前記対象ストレージ・ユニット・マッピング・テーブルを修正するステップであって、前記対象ストレージ・ユニット・マッピング・テーブル内の対応するポインタを前記ソース・データから前記修正空間へ転送するステップと、

少なくとも1つの対象ストレージ・ユニットを削除するステップであって、どの対象ストレージ・ユニットも前記ソース・ストレージ・ユニットを参照しない場合には、前記ソース・ストレージ・ユニットを通常のボリュームとしてマーク付けし、前記ソース・ストレージ・ユニットが通常のボリュームとしてマーク付けされた場合には修正可能である、ステップと

を含む、方法。

【請求項21】

ある期間に渡って、前記対象ストレージ・ユニット・マッピング・テーブルは、前記ソース・データの部分と、前記修正空間の部分とをポイントするポインタの固有の組み合わせを通じて、前記対象ストレージ・ユニットのための前記ソース・データの固有のコピーを保持する、請求項20に記載の方法。

【請求項22】

前記ポインタは、直接ポインタおよび間接ポインタのうちのいずれか1つを備える、請求項20に記載の方法。

【請求項23】

前記修正空間として利用可能なストレージの空きブロックのリストを保持するステップをさらに含む、請求項20に記載の方法。

【請求項24】

前記ソース・ストレージ・ユニット・マッピング・テーブルは、前記ソース・ストレージ・ユニット内の物理ストレージ装置上の位置に対するポインタを含む、請求項20に記載の方法。

【請求項25】

前記ソース・ストレージ・ユニット・マッピング・テーブルおよび前記対象ストレージ・ユニット・マッピング・テーブルは、それぞれ、

ポインタのマッピング・テーブル、

フラグのマッピング・テーブル、および

ハッシング・テーブルを伴うリンク付けされたリスト

10

20

30

40

50

のうちの 1 つを備える、請求項 20 に記載の方法。

【請求項 26】

ソース・ストレージ・ユニット内に記憶されたソース・データの仮想ポイント・イン・タイム・コピーを作成および保持するサービスであって、

前記ソース・データのポイント・イン・タイム・コピーを作成するための要求を受信するステップと、

前記ソース・データの前記コピーを作成する代わりに、前記ソース・データに対するポインタを含む対象ストレージ・ユニット・マッピング・テーブルを作成するステップと、

前記対象ストレージ・ユニットに関連付けられている修正空間であって、前記対象ストレージ・ユニットに固有の前記ソース・データに対する変更を記憶するだけである前記修正空間を保持するステップと、

前記修正空間に対してデータが書き込まれながら、前記対象ストレージ・ユニット・マッピング・テーブルを修正するステップであって、前記対象ストレージ・ユニット・マッピング・テーブル内の対応するポインタを前記ソース・データから前記修正空間へ転送するステップと

を含む、サービス。

【請求項 27】

ある期間に渡って、前記対象ストレージ・ユニット・マッピング・テーブルは、前記ソース・データの部分と、前記修正空間の部分とをポイントするポインタの固有の組み合わせを通じて、前記対象ストレージ・ユニットのための前記ソース・データの固有のコピーを保持する、請求項 26 に記載のサービス。

【請求項 28】

前記ポインタは、直接ポインタおよび間接ポインタのうちのいずれか 1 つを備える、請求項 26 に記載のサービス。

【請求項 29】

前記修正空間を設けるステップをさらに含む、請求項 26 に記載のサービス。

【請求項 30】

前記ソース・ストレージ・ユニット・マッピング・テーブルは、前記ソース・ストレージ・ユニット内の物理ストレージ装置上の位置に対するポインタを含む、請求項 26 に記載のサービス。

【請求項 31】

前記ソース・ストレージ・ユニット・マッピング・テーブルおよび前記対象ストレージ・ユニット・マッピング・テーブルは、それぞれ、

ポインタのマッピング・テーブル、

フラグのマッピング・テーブル、および

ハッシング・テーブルを伴うリンク付けされたリスト

のうちの 1 つを備える、請求項 26 に記載のサービス。

【請求項 32】

機械読み込み可能なプログラム・ストレージ装置であって、

ソース・ストレージ・ユニット内に記憶されたソース・データの仮想ポイント・イン・タイム・コピーを作成および保持する方法を実行するために、前記機械によって実行可能な命令のプログラムを具体的に実施するプログラム・ストレージ装置であって、前記方法は、

前記ソース・データの仮想ポイント・イン・タイム・コピーを作成するための少なくとも 1 つの要求を受信するステップと、

前記ソース・ストレージ・ユニットをソース・ボリュームとしてマーク付けするステップであって、前記ソース・ストレージ・ユニットがソース・ボリュームとしてマーク付けされる場合には修正はできない、ステップと、

前記ソース・データの前記コピーを作成する代わりに、前記ソース・データに対するポインタを含む対象ストレージ・ユニット・マッピング・テーブルを作成するステップと、

前記対象ストレージ・ユニットに関連付けられている修正空間であって、前記対象ストレージ・ユニットに固有の前記ソース・データに対する変更を記憶するだけである前記修正前記修正空間として利用可能なストレージの空きブロックのリスト空間を保持するステップと、

前記修正空間に対してデータが書き込まれながら、前記対象ストレージ・ユニット・マッピング・テーブルを修正するステップであって、前記対象ストレージ・ユニット・マッピング・テーブル内の対応するポインタを前記ソース・データから前記修正空間へ転送するステップとを含む、プログラム・ストレージ装置。

【請求項 3 3】

ある期間に渡って、前記対象ストレージ・ユニット・マッピング・テーブルは、前記ソース・データの部分と、前記修正空間の部分とをポイントするポインタの固有の組み合わせを通じて、前記対象ストレージ・ユニットのための前記ソース・データの固有のコピーを保持する、請求項 3 2 に記載のプログラム・ストレージ装置。

10

【請求項 3 4】

前記ポインタは、直接ポインタおよび間接ポインタのうちのいずれか 1 つを備える、請求項 3 2 に記載のプログラム・ストレージ装置。

【請求項 3 5】

ある期間に渡って、前記対象ストレージ・ユニット・マッピング・テーブルは、前記ソース・データの部分と、前記修正空間の部分とをポイントするポインタの固有の組み合わせを通じて、前記対象ストレージ・ユニットのための前記ソース・データの固有のコピーを保持する、請求項 3 2 に記載のプログラム・ストレージ装置。

20

【請求項 3 6】

前記ポインタは、直接ポインタおよび間接ポインタのうちのいずれか 1 つを備える、請求項 3 2 に記載のプログラム・ストレージ装置。

【請求項 3 7】

前記方法は、前記修正空間を設けるステップをさらに含む、請求項 3 2 に記載のプログラム・ストレージ装置。

【請求項 3 8】

前記ソース・ストレージ・ユニット・マッピング・テーブルは、前記ソース・ストレージ・ユニット内の物理ストレージ装置上の位置に対するポインタを含む、請求項 3 2 に記載のプログラム・ストレージ装置。

30

【請求項 3 9】

前記ソース・ストレージ・ユニット・マッピング・テーブルおよび前記対象ストレージ・ユニット・マッピング・テーブルは、それぞれ、  
ポインタのマッピング・テーブル、  
フラグのマッピング・テーブル、および  
ハッシング・テーブルを伴うリンク付けされたリスト  
のうちの 1 つを備える、請求項 3 2 に記載のプログラム・ストレージ装置。

【発明の詳細な説明】

【技術分野】

40

【0001】

本発明は、マイクロコンピュータ・システムのデータ・ストレージ機構に記憶されたコンピュータ・システム・データのポイント・イン・タイム・コピーを作成するための方法に関する。より特定的には、本発明は、対象となる 1 つ以上のボリュームについての空間的および時間的に効率のよいコピーを作成するための、最小限のリソースでかつソフトウェア・アルゴリズムを実行するのが容易な方法を説明する。

【背景技術】

【0002】

ポイント・イン・タイム・コピーを使用する従来のシステムは、元ディスクまたはソース・ディスクもしくは論理ユニット (LUN) を、物理的にデータをコピーすることなく

50

、対象LUN上のブロックがソースLUN上の対応するブロックを参照するような制御データ構成（メタデータ）を設定することによって、1つ以上のコピー・インスタンス（対象ディスクまたは対象LUNとも称される）にコピーする。そして、ソースLUN上の1つ以上のLUNが修正または書き込みもしくはその両方がなされた場合にのみ、これらのブロックをコピーする必要がある。従来のシステムにおいては、これらのブロックは、ソースLUN上で修正される前にすべてのコピー・インスタンスにコピーされる。これは、LUNのポイント・イン・タイム・コピーを作成できる数（例えば4つ）を限定することになる。この取り組みは、ブロックは修正されない場合には一箇所にのみ記憶される（空間的に効率のよいコピーと称される）という利点がある。性能関連の利点に加えて、これにより、キャッシュ性能を向上することができる。なぜならば、互いに異なるディスクまたはLUN上のディスク・ブロックは同一のディスク・ブロックを参照するので、キャッシュ・メモリ内のスペースは少なくすむだろうからである。他の取り組みにおいて、各対象LUNは、ソースLUNと同様のディスク空間を必要とし、ディスク・ブロックの大きな割合部分が読み出し専用の場合であっても、各コピー内に記憶される。ソースからターゲットLUNへのこのような物理コピーは、通常、バックグラウンド処理として行われる。この種のポイント・イン・タイム・コピー動作は、良好なコピーおよびキャッシング性能をもたらすことはない。

10

【発明の開示】

【発明が解決しようとする課題】

【0003】

20

したがって、ソース・データをコピーするための改良されたシステムおよび方法に対する必要性がある。

【課題を解決するための手段】

【0004】

本明細書における実施形態は、ソース・ストレージ・ユニットまたはソースLUN内に記憶されたソース・データの仮想ポイント・イン・タイム・コピーを作成および保持する方法およびサービスを含む。本方法／サービスは、ソース・データのポイント・イン・タイム・コピーを作成するための少なくとも1つの要求を受信する。しかしながら、ソース・データのコピーを作成する代わりに、本発明は、対象ストレージ・ユニットまたは同一のストレージ・システムによって管理される他のストレージ空間内に記憶される対象ストレージ・ユニットまたはLUNマッピング・テーブルを作成する。この対象ストレージ・ユニット・マッピング・テーブルは、ソース・データに対するポインタを含む。加えて、本発明は、ストレージ・システム内の修正空間を保持する。修正空間の各部分は、所定の対象LUNに関連付けられている。対象LUNのための修正空間は、対象LUNに固有のソース・データに対する変更を記憶するだけである。データが修正空間に書き込まれながら、対象ストレージ・ユニット・マッピング・テーブルは修正され、この修正は、対象ストレージ・ユニット・マッピング・テーブル内の対応するポインタをソース・データから修正空間へ転送することによって行われる。

30

【0005】

ソース・データの仮想コピーがある限り、本方法は、ソース・ストレージLUNをソース・ボリュームとしてマーク付けする。ソース・ストレージLUNがソース・ボリュームとしてマーク付けされる場合には、修正することはできない。代わりに、対象ストレージ・ユニット・マッピング・テーブル・ポインタおよび対応の修正のみ、変更することができる。対象ストレージ・ユニットは、時間の経過と共に削除することができる。ソース・ストレージ・ユニットを参照する対象ストレージ・ユニットがない場合には、本発明は、ソース・ストレージ・ユニットを通常のボリュームとしてマーク付けする。ソース・ストレージ・ユニットが通常のボリュームとしてマーク付けされる場合には、修正することができる。ソース・ボリュームを読み出し専用とすることによって、性能上なんら妥協することなく、数個のコピー・インスタンスを有することが可能になる。さらに、この要件により、このポイント・イン・タイム・コピー方法を容易に実施できることになる。なぜな

40

50

らば、各対象 L U N は、修正空間を管理する必要があるからである。

【 0 0 0 6 】

また、本明細書における実施形態は、ソース・データの仮想ポイント・イン・タイム・コピーを保持するためのコンピュータ・システムを含む。本システムは、（複数の）物理ストレージ装置を有するソース・ストレージ・ユニットまたは L U N と、ソース・ストレージ・マッピング・テーブルとを備える。ソース・ストレージ・マッピング・テーブルは、ソース・ストレージ・ユニット内のソース・データに対するポインタを保持し、ソース・ストレージ装置または同一のストレージシステムによって管理される他のストレージ装置に記憶される。（複数の）対象ストレージ・ユニットは、対象 L U N とも称され、これもシステムに含まれる。この対象ストレージ・ユニットは、ソース・ストレージ内のソース・データに対するポインタと、修正領域内に記憶された修正データに対するポインタとを保持する対象ストレージ・ユニット・マッピング・テーブルを含む。対象ストレージ・ユニット・マッピング・テーブル内のポインタが対応するソース・データの代わりに修正データをポイントする以外は、対象ストレージ・ユニット・マッピング・テーブル内のエントリは、ソース・ストレージ・ユニット・マッピング・テーブルないでポイントされたものと同じのデータブロックをポイントする。

10

【 0 0 0 7 】

ある期間に渡って、対象ストレージ・ユニット・マッピング・テーブルは、ソース・データの部分と、修正空間の部分とをポイントするポインタの固有の組み合わせを通じて、ソース・データの固有の仮想コピーを保持する。これらのポインタは、直接ポインタまたは間接ポインタを備えることができる。修正空間を使用して、修正データを保持する。これは、対象ストレージ・ユニットまたはこの目的専用のストレージ・システム上の特別な空間の一部である。ストレージ・システム内のストレージの空きブロックのリストは、修正空間として使用するのに利用可能なブロックを列挙する。ソース・ストレージ・ユニット・マッピング・テーブルは、ソース・ストレージ・ユニット内の物理ストレージ装置上の位置に対するポインタを含む。ソース・ストレージ・ユニット・マッピング・テーブルおよび対象ストレージ・ユニット・マッピング・テーブルは、ポインタのマッピング・テーブル、フラグのマッピング・テーブル、またはハッシング・テーブルを伴うリンク付けされたリストであってもよい。

20

【 0 0 0 8 】

ソース・データは、ソース・ストレージ・ユニットまたはソース L U N に保持される。ソース・ストレージ・ユニット・マッピング・テーブルは、対象ストレージ・ユニットまたは対象 L U N か、もしくは同一のストレージ・システムによって管理される他のストレージ・ユニットのいずれかに保持される。対象ストレージ・ユニット・マッピング・テーブルは、対象ストレージ・ユニットまたは対象 L U N か、もしくは同一のストレージ・システムによって管理される他のストレージ・ユニットのいずれかに保持される。修正空間は、対象ストレージ・ユニットまたは対象 L U N か、もしくは同一のストレージ・システムによって管理される他のストレージ・ユニットのいずれかの内部に保持される。

30

【 0 0 0 9 】

本発明のこれらおよび他の局面および目的は、以下の説明および添付の図面と共に考慮するといっそう認識および理解されるだろう。しかしながら、以下の説明は、本発明の好ましい実施形態およびその数多くの具体的な詳細を示しているものの、限定ではなく例示の目的で与えられていることが理解されるべきである。多くの変更および修正が、本発明の精神を逸脱することなく、本発明の範囲内で行われてもよい。本発明は、そのような修正のすべてを含むものである。

40

【 0 0 1 0 】

本発明は、図面を参照して以下の詳細の説明からいっそう理解されるだろう。

【 発明を実施するための最良の形態 】

【 0 0 1 1 】

本発明ならびにその様々な特徴および利点の詳細を、添付の図面に図示されかつ以下の

50



説明において詳述されている非限定的な実施形態を参照して、より完全に説明する。注意すべきなのは、図面に示された特徴は、必ずしも一律の縮尺で描かれているわけではないということである。周知の構成要素および処理手法の説明は、本発明を不必要に不明瞭にしないようにするために省略されている。本明細書において使用される例は、本発明が実施できるようなやり方についての理解を助け、さらに当業者が本発明を実施できるようにするために意図されているに過ぎない。従って、本例は、本発明の範囲を制限するものと解釈されるべきではない。

#### 【0012】

以下に説明する本発明は、ストレージ・コントローラ上でのポイント・イン・タイム・コピーを介するサーバ・プロビジョニングに特に適した、新しいコピー手法/サービス/システムを提供する。ソース・データのコピーを作成する代わりに、本発明は、対象ストレージ・ユニット・マッピング・テーブルを作成する。この対象ストレージ・ユニット・マッピング・テーブルは、ソース・データに対するポインタを含む。加えて、本発明は、ストレージ・システム内の修正空間を保持する。修正空間の各部分は、所定の対象ストレージ・ユニットまたは対象LUNに関連付けられている。修正空間は、対応の対象LUNに固有のソース・データに対する変更を記憶するだけである。データが修正空間に書き込まれながら、対象ストレージ・ユニット・マッピング・テーブルは修正され、この修正は、対象ストレージ・ユニット・マッピング・テーブル内の対応するポインタをソース・データから修正空間へ転送することによって行われる。ある期間に渡って、対象ストレージ・ユニット・マッピング・テーブルは、ソース・データの部分と、修正空間の部分とをポインタするポインタの固有の組み合わせを通じて、ソース・データの固有の仮想コピーを保持する。

10

20

#### 【0013】

本発明により、論理ディスクがコピー動作のソースとなる場合には、自動的に読み出し専用となることになる。この要件により、システムは、高性能を保持しつつ、より多くのコピーを提供することができる。さらに、元コピーが読み出し専用の場合、修正されていないすべてのディスク・ブロックは、ディスクに一度記憶され、互いに異なるコピーにアクセスする複数のノードによってアクセスされる場合には、1つのコピーのみがストレージ管理キャッシュに記憶されればよい。この動作は、オペレーティングシステムの画像を含むゴールデン・ディスクが作成されて複数(数十)のコピーが必要に応じて作成される場合のプロビジョニングに非常に適している。

30

#### 【0014】

論理ユニットまたはLUNと称される1つ以上のストレージ・ユニットを、空間的および時間的に効率のよいポイント・イン・タイム・コピー動作のソースとして使用することができる。この動作は、ソース・データブロックを他のストレージ領域に物理的にコピーする必要がないので、実行するのに長時間かからないため、時間的に効率がよい。この動作を行うのに必要な時間は、ユーザ要求をストレージ装置内の正確な位置に導くことが可能な比較的小さいデータ構成を作成および保持することに主に費やされる。また、本発明において提示されるこの動作は、空間的にも効率がよい。なぜならば、データブロックが新しい位置に物理的にコピーされるのは変更があった場合のみなので、ポイント・イン・タイム・コピーを維持するために使用される物理ストレージの量は最小限だからである。ソース・データのための空間は、基本的にはソースLUNのサイズであり、膨大なものでありうる。問題は、ソースのサイズに対する制御はないということである。対象に対しては、修正されたブロックがそれほど多くない限り、空間は非常に限られている。よって、対象のサイズは、修正されたブロックの量に左右されることになる。本開示は、1つのLUNのポイント・イン・タイム・コピーを作成するための実施形態を説明しているが、当業者によって理解されるように、同様の取り組みを、複数のLUNに対する動作を単一の動作として行うために使用することができる。

40

#### 【0015】

ソースLUNのポイント・イン・タイム・コピーが要求されると、本発明は、論理ユニ

50

ット毎に元ソース・データに対するポインタを保持するような、対象LUNのためのデータ構成（例えば、マッピング・テーブル）を設ける。一実施形態において、マッピング・テーブルは、ポインタおよび場合によっては他のフラグのテーブルとして実施される。他の実施形態において、マッピング・テーブルは、ハッシング・テーブルを伴うリンク付けされたリストとして実施される。

#### 【0016】

マッピング・テーブル内のエントリは、ソースLUNの対応するブロックの物理位置を直接ポイントするように、まず初期化される。間接ポインタが使用される他の実施形態において、ポインタは、ソースLUNのための同様のデータ構成内のエントリをポイントする。ソースLUNのためのエントリは、データ・ブロックが物理装置上で実際に記憶される場所についての情報を含む。

10

#### 【0017】

ある量のストレージ空間は、修正空間と称され、対象LUNの物理ストレージ装置内か、または各ポイント・イン・タイム・コピーのためのコピー動作を行うストレージ・システムによって管理される他のストレージ・ユニット内に取り置かれる。このストレージ空間のサイズは、一実施形態においては、コピー動作の作成者によって設定される。他の実施形態において、修正空間のサイズは、ストレージ・システム・ソフトウェアまたは管理者によって決定される。修正空間が使用されるのは、コピーLUNの論理ユニットまたはブロックが修正される場合である。この修正空間内の空きブロックのリストを含むデータ構成は、コピーLUN毎に維持される。

20

#### 【0018】

図1は、対象となるシステムを示し、ここにおいて、1つ以上のクライアント・マイクロプロセッサ・システム（110）がストレージ・システム（100）に接続されている。ストレージ・システムは、ストレージ・コントローラ（102）と称される制御ユニットと、場合によってはストレージ・システムの集積パッドであってもよいしそうでなくてもよい管理コンソール（104）と、ハード・ディスクなどの1つ以上の物理ストレージ装置（106）からなる。

#### 【0019】

図2は、1つのソースLUN（202）（ソースLUNまたはソース・ストレージ・ユニットとも称される）について3つのポイント・イン・タイム・コピー（210）（対象ボリュームまたはLUNとも称される）が作成されるシステムを示す。すべてのLUN（ソースLUNおよび対象LUNを含む）は、自身のマッピング・テーブル（204、212）を有する。ポイント・イン・タイム・コピーの初期化後、対象LUNのマッピング・テーブル内の各エントリは、ソース・ボリューム内の対応するエントリをポイントする。そのようなエントリの1つを図2に示す（206、214）。ソース・マッピング・テーブル内のブロック毎のエントリは、1つ以上のハード・ディスク・ドライブからなっているもよいような物理ストレージ（209）上のブロック（208）の位置をポイントする。注意すべきなのは、代わりに、対象マッピング・テーブルは、ソース・マッピング・テーブルの内容をコピーすることによって作成することもできるということである。そのような場合には、対象マッピング・テーブル内のエントリは、物理ストレージ上のブロックをポイントする直接ポインタである。

30

40

#### 【0020】

図3は、コピー・ボリュームにおける1つのブロックが修正される進歩的なシステムを示す。図示のように、コピー・マッピング・テーブル（310）内の対応するエントリは、ボリューム（312）の修正空間領域内の空き空間を使用してデータ・ブロック（314）を記憶する。

#### 【0021】

図4のフローチャート形式に示すように、ブロックに対する書き込みのための要求が受信されると、ストレージ・システムは、当該書き込みがポイント・イン・タイム・ソースLUN宛てかどうかを検出する（ブロック410）。もしそうであれば、アクセスは拒否

50

される（ブロック４２４）。なぜならば、本発明は、ソースＬＵＮが修正されることを許可しないからである。もしそうでなければ、ポイント・イン・タイム・コピーＬＵＮ宛てかどうかを知るために要求がチェックされる（ブロック４１２）。もしそうでなければ、これは、さらに検査することなく進めることができるアクセスである（ブロック４２６）。もしそうであれば、ストレージ・システムは、まず、マッピング・テーブルに照会する（ブロック４１４）。アクセスされようとするブロックについてのエントリが、当該ブロックがそのコピーのための修正空間に記憶されていることを示す場合には、新しいデータが対応の位置に書き込まれる（ブロック４２６）。対応のエントリが、当該ブロックがソースＬＵＮ内の元ブロックを通じてアクセスされようとしている場合には、修正空間にブロックを記憶するための場所を見つける必要がある。このためには、当該コピーのための空きブロックのリストに照会して、必要な空き空間が利用可能かどうかを調べる（ブロック４１６）。書き込み要求に対応するための空間がない場合には、書き込み要求は拒否されて、適切な動作が取られる（ブロック４２４）。例えば、空間の欠如を管理者に知らせるメッセージをシステム管理者に対して送ることができる。他の実施形態において、当該コピーは使用不可能であるとマーク付けされて、当該コピーに対するすべてのさらなるアクセスは拒否される。さらに他の実施形態において、コピー・ボリュームに対するアクセスは、新しいストレージ空間を必要としない限り可能とされる。書き込み要求に対応するための空き空間が利用可能である場合には、必要な（複数の）ブロックが空きリストから取り出されて（ブロック４１８）、データがこれらのブロックに記憶され（ブロック４２０）、マッピング・テーブルは、当該変更を反映するように更新される（ブロック４２２）。

10

20

#### 【００２２】

本発明は、ポイント・イン・タイム・コピー動作の実施を改良するものである。コピー毎に専用の空間（本原稿においては修正空間と称する）を設けることによって、ストレージ空間を、各ボリュームのブロックが任意の利用可能な位置に記憶できるようなグローバル・エンティティとして管理する必要がなくなる。さらに、ソース・ボリュームを読み出し専用ボリュームとすることによって、本願のコピー動作の拡張性（scalability）が大幅に向上する。この限定が実施されないと、ソース・ボリューム内のブロックが修正される場合に、ストレージ・システムは、コピー・ボリューム毎に修正空間内の空き空間を探して、元ブロックを１つずつそこにコピーして、コピー・ボリュームすべてについてマッピング・テーブルを更新する必要がある。この要件は、いかなるソース・ボリュームもサポート可能なコピー・ボリュームの数に対して上限を制限する。ソースから第１のコピー・ボリュームが作成されるとすぐにソース・ボリュームを読み出し専用であるとマーク付けすることによって、そのような要件をなくす。図４に示すように、ソースまたはコピー・ボリュームに対する各アクセスに対する評価および、もし許可されていればその実行は、非常に簡単に行うことができる。

30

40

#### 【００２３】

コピー・ボリュームは削除することができる。各ソース・ボリュームは、サポートするコピーの数のカウントを取る。ソースのコピーが削除される度に、コピー・カウンタがデクリメントされる。カウントがゼロに達すると、ソース・ボリュームは通常のボリュームとしてマーク付けされて、もはやソース・ボリュームではなくなる。これにより、ソース・ボリュームに対して適用された読み出し専用という制限はなくなり、ボリュームの内容は修正可能となる。コピーがボリュームから作成されるとすぐに、当該ボリュームはソース・ボリュームとマーク付けされて、コピー・カウンタの値が、第１のコピー用に１に設定され、新しいコピーが作成される度にインクリメントされる。

#### 【００２４】

よって、上述のように、本発明は、数多くの環境において有用なストレージ・ボリュームのポイント・イン・タイム・コピーを作成するための新しい取り組みを提供する。特に、本発明は、プロビジョニング環境と共に使用されるのが適している。そのような環境において、システム・ユーザによって使用されるオペレーティング・システムおよびアプリ

50

ケーションを含む1つ以上のゴールデン画像が作成される。ユーザの数および種類によっては、本システムは、時間的および空間的に効率のよいやり方で複数のコピーを作成するために使用されるゴールデン画像をロードする。そして、これらのコピーは、プロビジョニング・サーバのために使用される。

【0025】

本発明を好ましい実施形態の観点から説明してきたが、本発明は添付の請求項の精神および範囲内の修正を伴って実施可能であることを当業者は認識するだろう。

【図面の簡単な説明】

【0026】

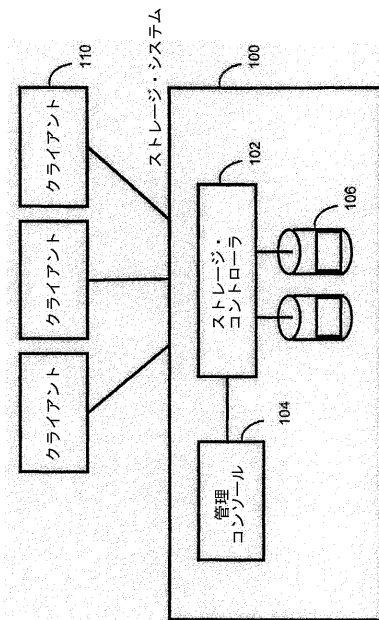
【図1】図1は、本発明の対象となるシステムの例を示す概略図である。

【図2】1つのボリュームについて3つのポイント・イン・タイム・コピーを伴うシステムの一例を示す概略図である。

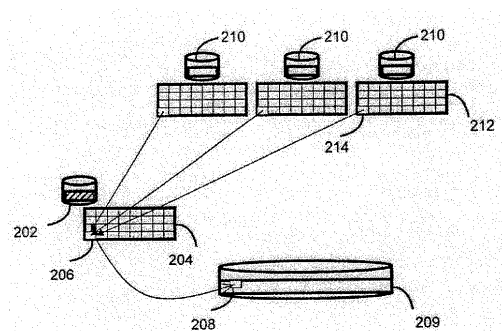
【図3】1つのコピー・ボリュームにおける1つのブロックが書き込まれた後の分離された空き空間管理を伴うシステムを示す概略図である。

【図4】本発明の一実施形態のフロー図を示す。

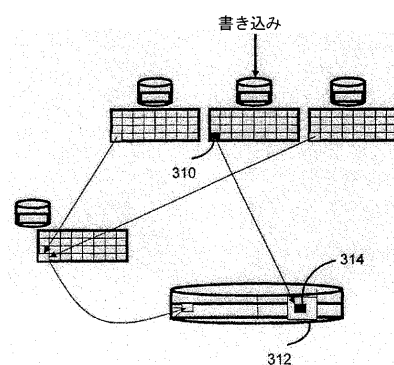
【図1】



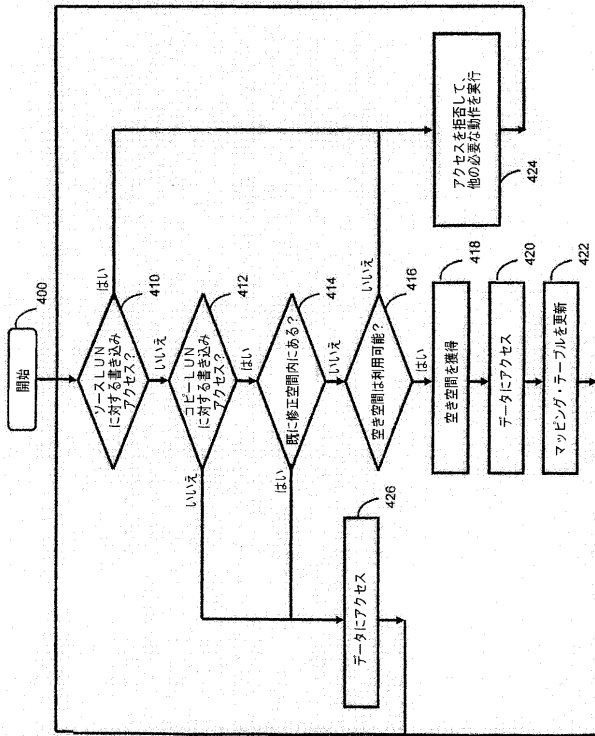
【図2】



【図3】



【図 4】



## 【国際調査報告】

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US2005/030847

<b>A. CLASSIFICATION OF SUBJECT MATTER</b> INV. G06F3/06 G06F11/14		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b> Minimum documentation searched (classification system followed by classification symbols) G06F		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the International search (name of data base and, where practical, search terms used) EPO-Internal, WPI Data, COMPENDEX		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	US 2002/112022 A1 (KAZAR MICHAEL L ET AL) 15 August 2002 (2002-08-15)  figures 2-4,6 paragraph [0007] paragraph [0067] - paragraph [0074] paragraph [0079] - paragraph [0081] paragraph [0089] - paragraph [0091] paragraph [0099] ----- -/--	1-13  20-25, 32-39
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents : *A* document defining the general state of the art which is not considered to be of particular relevance *E* earlier document but published on or after the international filing date *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another claim or other special reason (as specified) *O* document referring to an oral disclosure, use, exhibition or other means *P* document published prior to the international filing date but later than the priority date claimed *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. *Z* document member of the same patent family		
Date of the actual completion of the International search  9 August 2006		Date of mailing of the international search report  22/08/2006
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016		Authorized officer  Andlauer, J.C.

5

Form PCT/ISA/210 (second sheet) (April 2005)

page 1 of 2

## INTERNATIONAL SEARCH REPORT

International application No  
PCT/US2005/030847

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	US 2002/178335 A1 (SELKIRK STEPHEN S ET AL) 28 November 2002 (2002-11-28)  figures 7-10, 18C, 20A, 20B paragraphs [0019] - [0022] paragraph [0030] paragraph [0032] paragraphs [0081] - [0093] paragraph [0171] - paragraph [0177]	14-19, 26-31, 20-25, 32-39
A	"Compressed Dasd Emulation" [Online] 12 July 2004 (2004-07-12), , XP002393719 Retrieved from the Internet: URL: <a href="http://web.archive.org/web/20040712050939/http://www.conmicro.cx/hercules/cckdda_sd.html">http://web.archive.org/web/20040712050939/http://www.conmicro.cx/hercules/cckdda_sd.html</a> > [retrieved on 2006-08-01] the whole document	1-39
A	US 2003/195903 A1 (MANLEY STEPHEN L ET AL) 16 October 2003 (2003-10-16) figures 5-7 paragraphs [0070] - [0072] paragraphs [0033] - [0035]	1-39

5

Form PCT/ISA/210 (continuation of second sheet) (April 2005)

page 2 of 2

## INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No  
PCT/US2005/030847

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2002112022 A1	15-08-2002	AU 2740502 A EP 1358582 A1 WO 0250433 A1 US 2005033748 A1	01-07-2002 05-11-2003 27-06-2002 10-02-2005
US 2002178335 A1	28-11-2002	NONE	
US 2003195903 A1	16-10-2003	US 2003182325 A1	25-09-2003



## フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW

(74)代理人 100086243

弁理士 坂口 博

(72)発明者 バニカゼミ、モハメッド

アメリカ合衆国 1 0 0 2 3 ニューヨーク州 ニューヨーク ブロードウェイ 2 0 2 5 1 2 C

(72)発明者 ポフ、ダン

アメリカ合衆国 1 0 5 4 1 ニューヨーク州 マホバック リッジ・レーン 3

Fターム(参考) 5B065 BA01 CA13 EA33

## 【要約の続き】

き込まれながら、対象ストレージ・ユニット・マッピング・テーブル内の対応するポインタをソース・データから修正空間へ転送することによって修正される。

【選択図】図 1