

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第4965456号  
(P4965456)

(45) 発行日 平成24年7月4日(2012.7.4)

(24) 登録日 平成24年4月6日(2012.4.6)

(51) Int.Cl.

F 1

G06F 12/00 (2006.01)  
G06F 3/06 (2006.01)G06F 12/00 514A  
G06F 3/06 301S

請求項の数 20 (全 12 頁)

(21) 出願番号 特願2007-544873 (P2007-544873)  
 (86) (22) 出願日 平成17年11月17日 (2005.11.17)  
 (65) 公表番号 特表2008-523485 (P2008-523485A)  
 (43) 公表日 平成20年7月3日 (2008.7.3)  
 (86) 國際出願番号 PCT/EP2005/056040  
 (87) 國際公開番号 WO2006/061316  
 (87) 國際公開日 平成18年6月15日 (2006.6.15)  
 審査請求日 平成20年9月12日 (2008.9.12)  
 (31) 優先権主張番号 11/009,937  
 (32) 優先日 平成16年12月10日 (2004.12.10)  
 (33) 優先権主張国 米国(US)

(73) 特許権者 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】システムと共有バッファ内のストレージとの間でデータを転送する方法、システム、及びコンピュータ・プログラム(システムと共有バッファ内ストレージとの間のデータ転送)

## (57) 【特許請求の範囲】

## 【請求項 1】

アプリケーションが第1のコンポーネントからバッファを要求するステップと、  
 第1のコンポーネントがバッファを割り振るステップと、  
 第1のコンポーネントがアプリケーション・データに関するバッファの第1のオフセットと、第1のヘッダに関するバッファの第2のオフセットとをアプリケーションに返すステップであって、前記第1のオフセットと前記第2のオフセットは、それ以前に書き込まれたデータの前に第1のヘッダと第2のヘッダを書き込むのに十分な量のスペースが設けられるように計算される、前記返すステップと、  
 アプリケーションがアプリケーション・データをバッファの第1のオフセットに書き込むステップと、  
 アプリケーションが第1のヘッダをバッファ内の第2のオフセットに書き込むステップと、  
 第1のコンポーネントが第2のヘッダをバッファの第3のオフセットに書き込むステップと、  
 第1のコンポーネントが第2のヘッダを指すポインタを第2のコンポーネントに受け渡すステップと、  
 第2のコンポーネントがポインタの受け取りに応答して、バッファ内容をどう処理するかを決定するよう、前記バッファ内の第2のヘッダを処理するステップと、  
 を有する方法。

**【請求項 2】**

第1のオフセットは第2のオフセットよりも大きく、第2のオフセットは第3のオフセットよりも大きい、請求項1に記載の方法。

**【請求項 3】**

第1のコンポーネントが第3のヘッダをバッファの第4のオフセットに書き込むステップを更に有し、第1のオフセットは第2のオフセットよりも大きく、第2のオフセットは第3のオフセットよりも大きく、第3のオフセットは第4のオフセットよりも大きい、  
請求項1に記載の方法。

**【請求項 4】**

第2のコンポーネントが第2のヘッダをバッファ内容をストレージ・デバイスに転送するのに使用される情報に置き換えるステップ、  
を更に有する請求項1に記載の方法。

10

**【請求項 5】**

第3のオフセットにおける第2のヘッダはアプリケーション・データ及びヘッダを有するバッファ内容を第1のデバイス又は第2のデバイスに転送するよう指示し、

第1のコンポーネントが第2のヘッダを指すポインタを第2のコンポーネントに受け渡すステップと、

第2のコンポーネントが、バッファ内容を第1のデバイスと第2のデバイスのどちらに送信すべきか判定するために第2のヘッダを処理するステップと、

前記第2のヘッダが前記バッファ内容を前記第1のデバイスに送信するよう指示していることが判定されたことに応じて、前記バッファ内容を前記第1のデバイスに転送するのに使用されるデバイス情報を第2のヘッダに書き込むステップであって、デバイス情報が第1のコンポーネントによって書き込まれた第2のヘッダ情報と置き換えられるようにしたステップと、

20

を更に有する請求項1に記載の方法。

**【請求項 6】**

第1のデバイスはストレージ・デバイスからなり、  
バッファ内容をストレージ・デバイスに送信すべきことが判定されたことに応じて、ストレージ・デバイス用のストレージ・デバイス・ドライバが第2のヘッダのデバイス情報を使用してバッファ内容をバッファからストレージ・デバイスに転送するステップ、  
を更に有する請求項5に記載の方法。

30

**【請求項 7】**

前記第1のコンポーネントから受け渡される前記ポインタは第1のポインタからなり、  
バッファ内容をストレージ・デバイスに送信すべきことが判定されたことに応じて、第2のコンポーネントがバッファを指す第2のポインタを第3のコンポーネントに送り、第3のコンポーネントがデバイス情報を前記第2のヘッダに書き込むステップと、

第3のコンポーネントがバッファを指す第3のポインタをストレージ・デバイス・ドライバに送り、ストレージ・デバイス・ドライバが第3のポインタを使用してバッファの内容をストレージ・デバイスのストレージ・メディアに書き込むステップと、

を更に有する請求項6に記載の方法。

40

**【請求項 8】**

前記第1のコンポーネントは、ストレージ・マネージャのアプリケーション・プログラム・インターフェイス(API)からなり、第3のコンポーネントはストレージ・エージェントからなる、請求項7に記載の方法。

**【請求項 9】**

前記第2のデバイスはストレージ・サーバからなり、  
第2のヘッダがバッファ内容をストレージ・サーバに送信するよう指示していることが判定されたことに応じて、第2のコンポーネントがバッファ内容をストレージ・サーバに送信するステップ、  
を更に有する請求項5に記載の方法。

50

**【請求項 10】**

前記第1のコンポーネントが、第1のオフセットを、アプリケーション・データを第1のオフセットから書き込んだ後に第1及び第2のヘッダを書き込むのに十分な量のスペースがバッファの第1のオフセットの前に設けられるように、計算するステップと、

前記第1のコンポーネントが、第2のオフセットを、アプリケーション・データ及び第1のヘッダを書き込んだ後に第2のヘッダを書き込むのに十分な量のスペースがバッファの第2のオフセットの前に設けられるように、計算するステップと、

を更に有する請求項1に記載の方法。

**【請求項 11】**

前記第1のコンポーネントが、第3のオフセットを、アプリケーション・データ及び第1のヘッダを書き込んだ後に第2のヘッダを書き込むのに十分な量のスペースがバッファの第3のオフセットの前に設けられるように、計算するステップ、10

を更に有する請求項10に記載の方法。

**【請求項 12】**

アプリケーション・データが固定の数の倍数であるかどうかを判定するステップを更に含み、第1及び第2のオフセットは、アプリケーション・データが固定の数の倍数でないことが判定されたことに応じて、空きスペース及びアプリケーション・データが固定の数の倍数を占有するように、第1のオフセットに書き込まれたアプリケーション・データの前に空きスペースを確保するように計算され、第1のヘッダは第2のオフセットから空きスペースの開始位置まで延びる、20

請求項1に記載の方法。

**【請求項 13】**

第1のヘッダが固定の数の倍数であるかどうかを判定するステップを更に含み、第3のオフセットは、第1のヘッダが固定の数の倍数でないことが判定されたことに応じて、第1のヘッダ及び空きスペースが固定の数の倍数を占有するように、第1のヘッダの前に空きスペースを確保するように選択され、第2のヘッダは、第3のオフセットから第1のヘッダの前の空きスペースの開始位置まで延びる、

請求項12に記載の方法。

**【請求項 14】**

前記固定の数は4キロバイトを有し、アプリケーション・データが固定の数の倍数でなくなるのは、バッファの第1のオフセットに書き込まれたアプリケーション・データがバッファの4キロバイト境界上に位置合わせされない場合である、請求項12に記載の方法30。  
。

**【請求項 15】**

アプリケーションが複数のバッファを要求するステップと、

第1のコンポーネントが複数のバッファをアプリケーションに返すステップと、

アプリケーションがアプリケーション・データと第1のヘッダとを複数のバッファに同時に書き込むステップと、

を更に有する請求項1に記載の方法。

**【請求項 16】**

第1のコンポーネントが第2のコンポーネントからバッファを要求するステップと、40

第2のコンポーネントがバッファを割り振るステップと、

第2のコンポーネントがアプリケーション・データに関するバッファの第1のオフセットと、第1のヘッダに関するバッファの第2のオフセットとを第1のコンポーネントに返すステップであって、前記第1のオフセットと前記第2のオフセットは、それ以前に書き込まれたデータの前に第1のヘッダと第2のヘッダを書き込むのに十分な量のスペースが設けられるように計算される、前記返すステップと、

第2のコンポーネントがアプリケーション・データをバッファの第1のオフセットに書き込むステップと、

第2のコンポーネントが第1のヘッダをバッファ内の第2のオフセットに書き込むステ50

ップと、

第3のコンポーネントが第2のヘッダをバッファ内の第3のオフセットに書き込むステップと、

を有する方法。

【請求項17】

第1のオフセットは第2のオフセットよりも大きく、第2のオフセットは第3のオフセットよりも大きい、

請求項16に記載の方法。

【請求項18】

第1のコンポーネントがストレージ・デバイスから得られたアプリケーション・データをバッファに書き込むステップと、

10

第2のコンポーネントが、バッファへのポインタを、前記アプリケーション・データについての要求を開始したアプリケーションへ受け渡すステップと、

を更に有する請求項16に記載の方法。

【請求項19】

プロセッサと、

プロセッサからアクセス可能であり、複数のバッファを有するメモリと、

請求項1乃至18のいずれかの請求項に記載の方法のステップをプロセッサに実行させるためのプログラムと、

を備えるシステム。

20

【請求項20】

それぞれ複数のバッファを含むメモリと通信し、請求項1乃至18のいずれかの請求項に記載の方法のステップをコンピュータに実行させるためのコンピュータ・プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、システムと共有バッファのストレージとの間でデータを転送する方法、システム、及びプログラムに関するものである。

【背景技術】

【0002】

30

ネットワーク・バックアップ環境において、クライアント・システムは、データをストレージ・デバイスにバックアップし、バックアップをバックアップ・サーバと協調させることができる。例えば、インターナショナル・ビジネス・マシーンズ(「IBM」(R))社の製品であるTivoli(R)ストレージ・マネージャは、クライアント・サーバ・システムがクライアント・データをバックアップするためのソフトウェアを提供する(IBM及びTivoliはIBM Corporationの登録商標である)。クライアント・マシンは、クライアント・アプリケーションとストレージ・デバイスとの間で転送されるデータを処理する複数のプログラム・コンポーネントを含むことができる。各コンポーネントは、以前のコンポーネントによって使用されたデータをバッファから読み出し、次に、データをコンポーネントに割り振られたバッファにコピーするメモリ・コピー操作を実施する。データをコンポーネントに割り振られたバッファの間で受け渡されるデータとしてのデータの処理の一環として、プログラム・コンポーネントはヘッダ情報をデータに追加することができる。プログラム・コンポーネントは、データを以前のコンポーネントのバッファから読み出し、データを処理し、そして新しいヘッダをデータに追加した後、このデータ及び新しいヘッダをそのコンポーネントに割り当てられたバッファに書き込む。このデータを処理しているコンポーネント・チェーンの次のプログラム・コンポーネントもこれと同様の処理を行うことができる。

40

【発明の開示】

【発明が解決しようとする課題】

【0003】

50

システムと共有バッファ内のストレージとの間でデータを転送する方法、システム、及びプログラムを提供する。

【課題を解決するための手段】

【0004】

本発明の第1の態様において、アプリケーションはコンポーネントからバッファを要求する。コンポーネントはバッファをアプリケーションに割り振り、アプリケーション・データに関するバッファの第1のオフセットと、第1のヘッダに関するバッファの第2のオフセットとをアプリケーションに返す。アプリケーションはアプリケーション・データをバッファの第1のオフセットに書き込み、第1のヘッダをバッファの第2のオフセットに書き込む。第2のヘッダはバッファの第3のオフセットに書き込まれる。

10

【0005】

本発明の第2の態様において、第1のコンポーネントは第2のコンポーネントからバッファを要求する。第2のコンポーネントはバッファを第1のコンポーネントに割り振り、アプリケーション・データに関するバッファの第1のオフセットと、第1のヘッダに関するバッファの第2のオフセットとを第1のコンポーネントに返す。第2のコンポーネントはアプリケーション・データをバッファの第1のオフセットに書き込み、第1のヘッダをバッファ内の第2のオフセットに書き込む。第2のヘッダはバッファの第3のオフセットに書き込まれる。

【0006】

ここで、単なる例示のために、本発明の諸実施形態について添付図面を参照しながら説明する。

20

【発明を実施するための最良の形態】

【0007】

以下の記載では、本明細書の一部を形成し、本発明のいくつかの実施形態を例示する添付図面を参照する。本発明の範囲から逸脱することがなければ、他の実施形態が利用されてもよく、構造上及び動作上の変更が施されてもよいことが理解されるだろう。

【0008】

図1は、メモリ4を含むストレージ・クライアント2を有するネットワーク・コンピューティング環境を示している。ストレージ・クライアント2は、プログラム・コンポーネントとして、データベース・アプリケーション等のようなアプリケーション6と、ストレージ・マネージャのアプリケーション・プログラム・インターフェイス( API )8と、仮想サーバ10と、ストレージ・エージェント12と、ストレージ・デバイス16に対する書き込みインターフェイスを提供するストレージ・デバイス・ドライバ14とを含んでいる。ストレージ・マネージャAPI8はバッファ・プール20内のバッファ18を管理する。

30

【0009】

一実施形態において、ストレージ・デバイス・ドライバ14は、ストレージ・デバイス16に対する書き込みデータをストレージ・エリア・ネットワーク( SAN )等のようなネットワーク22を介してストレージ・デバイス16に通信する。一代替実施形態において、クライアント2は周辺機器相互接続( PCI )バスや小型コンピュータ・システム・インターフェイス( SCSI )バス等のようなバス・インターフェイスを介してストレージ・デバイス16に接続することができる。

40

【0010】

一実施形態において、仮想サーバ10はストレージ・デバイス16に対する書き込みデータ情報をローカル・エリア・ネットワーク( LAN )等のような第2のネットワーク28を介してストレージ・マネージャ・サーバ26を有するサーバ24に通信することができる。このようにして、ストレージとの間でやりとりされるデータは、SAN等のような第1のネットワークを介して伝送され、転送中のデータに関する制御情報は、LAN等のような別のネットワークを介して送られる。一代替実施形態では、データ及び情報とデータ転送処理に関するコマンドの両方を、図1に示されるような別個のネットワークではな

50

く同一のネットワークを介して伝送することができる。サーバ26は、ストレージ・デバイス16に記憶されているデータに関する情報をバックアップ・データベースで維持し、バックアップ・データ・セットを管理する。

#### 【0011】

ストレージ・マネージャAPI8は、アプリケーション6とストレージ・デバイス16との間で転送されるデータを記憶して使用するために、バッファ・プール20のバッファ18をクライアント2の各コンポーネントに割り振る。図2は、コンポーネント6、8、10、12及び14がバッファ18に追加する、アプリケーション6によって書き込まれるアプリケーション・データ40及びアプリケーション・ヘッダ44、ストレージ・マネージャAPI8によって書き込まれるAPIヘッダ46およびストレージ・マネージャAPI8によって書き込まれたバッファをどのように処理するかを指示するverb情報48を含む情報を示している。一実施形態において、ストレージ・エージェント12は、デバイス固有情報をもつverb情報を上書きして、ストレージ・デバイス・ドライバ14がデータをストレージ・デバイス16に転送できるようにすることができる。本明細書に記載の諸実施形態では、コンポーネント6、8、10、12及び14等のコンポーネントは、バッファ内容を処理するために、コンポーネントがバッファに書き込んだヘッダ又は他の情報を含むバッファの内容を指すポインタをコンポーネント8、10、12、及び14等の次のコンポーネントに伝達する。このようにして、バッファ18は各コンポーネントによって共有され、したがって、バッファのデータを処理しヘッダ情報を追加する各コンポーネントは、バッファの内容をコピーする追加的なバッファを割り振る必要がない。その代わりに、各コンポーネントはデータを処理し、ヘッダ及び他の情報を同一の共有バッファに追加することができ、それによって別個のバッファとの間のメモリ・コピー操作を実施する必要が解消される。

#### 【0012】

図3、図4及び図5は、データをストレージ・デバイス16に転送するためにデータをコンポーネント6、8、10、12及び14等のコンポーネント間で転送するためにコンポーネントで実施する処理を示している。図3を参照すると、アプリケーション6はストレージ・マネージャAPI8に1つ又は複数のバッファ18を要求することによって(ブロック102)アプリケーション・データをストレージ・デバイス16に転送する処理を開始する(ブロック100)。この要求はアプリケーション6が書き込もうとするヘッダの最大の長さを含むことができ、この長さはアプリケーション6がヘッダを書き込まない場合にはゼロとすることもできる。それに応じて、ストレージ・マネージャAPI8は、バッファ・プール20内のフリー・バッファ18として指示された1つ又は複数のバッファ18をアプリケーションに割り振る(ブロック104)。第1のコンポーネント又はストレージ・マネージャAPI8は、アプリケーション・データ40に関する第1のオフセットを、アプリケーション・データ40を第1のオフセットから書き込んだ後に第1及び第2のヘッダ例えばアプリケーション・ヘッダ44及びAPIヘッダ46を書き込むのに十分な量のスペースがバッファ18内の第1のオフセットの前に設けられるように、計算する(ブロック106)。第1のオフセットは要求で指示されたヘッダの長さに対するバッファ18内の十分なスペースを提供することができる。かくして、第1のオフセットはアプリケーション・データ40のバッファ18内の開始アドレスを指定する。第1のコンポーネント8は更に、第2のオフセットを、アプリケーション・データ40及びアプリケーション・ヘッダ44を書き込んだ後に第2のヘッダ例えばAPIヘッダ46を書き込むのに十分な量のスペースがバッファ内第2のオフセットの前に設けられるように、計算する(ブロック108)。このようにして、各オフセットは、一つのコンポーネントがそれ以前のコンポーネントによって書き込まれたデータの前にデータを書き込むことが可能となるように計算され、それによって各コンポーネントはそれ以前に書き込まれたアプリケーション・データ40及びヘッダの前にヘッダを追加することが可能となる。更に、バッファ18内のオフセット及び各ヘッダの長さを含めて、ヘッダ48、46、44等のすべてのヘッダ及びデータ40に関する情報を有するデータ構造が、バッファ18のトップ

10

20

30

40

50

に存在するようにし、コンポーネント 6、8、10、12 及び 14 のいずれもがバッファ 18 内のヘッダ及びデータの位置を判定することが可能になる。

#### 【0013】

本明細書に記載の諸実施形態において、ストレージ・マネージャ API 8 は、データが 4 K 境界等の固定された境界領域上に確実に留まることを望むこともある。アプリケーション 6 がストレージ・デバイス 16 との間の I/O を実行するときは、バッファ 18 が I/O 用のカーネルに受け渡される。当該カーネルは、ダイレクト・メモリ・アドレッシング (DMA) を実現するために、I/O デバイス (PCI バス、SCSI バス、SAS 等) がデータにアクセスし、当該データを適切なデバイスに送信することができるよう、バッファ 18 を (仮想メモリではなく) 物理メモリ 4 に「ピン (pin)」することができる。このようなバッファ 18 の「ピニング (pinning)」は、I/O 転送の処理中に (I/O 転送がバス上にある段階で) データが他のアプリケーションによって使用されないようにするものである。10

#### 【0014】

ポインタのアプリケーション・データ 40 が固定の数 (例えば 4 キロバイト (Kb)) の倍数でない場合には (ブロック 110)、第 1 のコンポーネント (ストレージ・マネージャ API 8) は、第 1 及び第 2 のオフセットを、空きスペース及びアプリケーション・データが固定の数、即ち 4 Kb の倍数を占有するように、第 1 のオフセットに書き込まれたアプリケーション・データの前に空きスペースを確保するように、計算する (ブロック 112)。例えば、第 2 のオフセットは、アプリケーション・データ 40 を 4 Kb 境界上に位置合わせするのに必要な空きスペース 42 の量と等しい、第 1 のオフセットからのオフセットを有することができ、それによって、書き込まれるヘッダの長さが第 2 のオフセット、例えばアプリケーション・ヘッダ 44 から始まるようにすることができる。第 1 のコンポーネント 8 は、第 1 及び第 2 のオフセットと、それ自体に割り振られているバッファ 18 のハンドルとをアプリケーション 6 に返す (ブロック 114)。各オフセットに応じて、アプリケーション 6 は、アプリケーション・データ 40 をバッファ 18 内の第 1 のオフセットに書き込み (ブロック 116)、第 1 のヘッダ、例えばアプリケーション・ヘッダ 44 をバッファ 18 内の第 2 のオフセットに書き込む (ブロック 118)。アプリケーション・ヘッダ 44 の書き込みは任意選択としてもよい。書き込みを行った後、アプリケーション 6 は、上記のバッファのハンドルを第 1 のコンポーネント (ストレージ・マネージャ API 8) に受け渡す (ブロック 120)。2030

#### 【0015】

図 4 を参照すると、第 1 のコンポーネント (ストレージ・マネージャ API 8) は、上記のバッファのハンドルをアプリケーション 6 から受け取ったときは (ブロック 130)、第 3 のオフセットを、アプリケーション・データ 40 及び第 1 のヘッダ (アプリケーション・ヘッダ 44) を書き込んだ後に第 2 のヘッダ (API ヘッダ 46) を書き込むのに十分な量のスペースがバッファ内の中 3 のオフセットの前に設けられるように、計算する (ブロック 132)。一実施形態において、(ブロック 134 で) 第 1 のヘッダ (アプリケーション・ヘッダ 44) が固定の数、例えば 4 Kb の倍数である場合には、ストレージ・マネージャ API 8 は、第 3 のオフセットを、第 1 のヘッダ及び空きスペースが固定の数 (4 Kb) の倍数を占有するように、第 1 のヘッダの前に空きスペースを確保するように、選択する (ブロック 136)。第 2 のヘッダ (API ヘッダ 46) は、バッファ 18 内の第 3 のオフセットに書き込まれる (ブロック 138)。ストレージ・マネージャ API 8 は更に、バッファ内容をストレージ・デバイス 16 とサーバ 24 のどちらに転送すべきか指示する第 3 のヘッダ (verb ヘッダ 48) をバッファ 18 内の第 4 のオフセットに書き込む (ブロック 140)。ストレージ・マネージャ API 8 は、ストレージ・デバイス 16 宛てのデータについてはストレージ・デバイス 16 に転送し、バックアップ・コマンド及び関連情報についてはクライアント 2 のバックアップ処理を管理するサーバ 24 に転送するよう指示する。次に、ストレージ・マネージャ API 8 は、第 3 のヘッダを指すポインタを仮想サーバ 10 等のような第 2 のコンポーネントに受け渡す (ブロック 144)。4050

42)。

#### 【0016】

仮想サーバ10は、第3のヘッダを指すポインタを受け取ったことに応じて、第3のヘッダ、例えばverbヘッダ48がバッファ内容をストレージ・デバイス16とサーバ24のどちらに送信するよう指示しているか判定する(ブロック144)。(ブロック144で)バッファ内容がサーバ24宛てのものである場合には、仮想サーバ10は、当該バッファ内容をネットワーク28を介してサーバ24に送信する(ブロック146)。そうではなく、(ブロック144で)バッファ内容がストレージ・デバイス16宛てのものである場合には、仮想サーバ10は、バッファ18を指す第2のポインタをストレージ・エージェント12等のような第3のコンポーネントに送る(ブロック148)。

10

#### 【0017】

図5を参照すると、第3のコンポーネント(ストレージ・エージェント12)は、第2のポインタを受け取ったことに応じて(ブロック150)、バッファ内容を第1のデバイスに転送するのに使用されるデバイス情報を、第3のヘッダに書き込む(ブロック152)。このデバイス固有情報は、第1のコンポーネント(ストレージ・マネージャAPI8)によって書き込まれた第2のヘッダ情報と置き換える。次に、ストレージ・エージェント12は、バッファ18を指す第3のポインタをストレージ・デバイス・ドライバ14に送る(ブロック154)。それに応じて、ストレージ・デバイス・ドライバ14は、第3のポインタを使用してバッファ18の内容にアクセスし、当該内容をストレージ・デバイス16内の磁気記憶媒体(例えばハード・ディスク・ドライブや磁気テープ等)、光記憶媒体、あるいは当業界で周知の他の適切な不揮発性記憶媒体等のようなストレージ・メディアに転送する(ブロック156)。

20

#### 【0018】

他の実施形態では、上記のヘッダのいくつかは書き込まれないこともある。例えば、アプリケーション6又はストレージ・マネージャAPI8がヘッダ44又は46を書き込まなかつた場合、第2のヘッダはverbデータ48からなることになる。別法として、各コンポーネントは、更なるオフセット計算が必要となる追加的なヘッダを書き込むことができる。更に、任意の書き込み済みのヘッダのオフセットは、それ以前に書き込まれたヘッダに追加すべき空きスペースを考慮に入れて、当該空きスペース及び書き込み済みのヘッダが4Kb境界上にくることを保証することもできる。このようにして、空きスペース42の後のオフセットに書き込まれるヘッダ44は、当該オフセットから空きスペース42の開始位置まで延びる。

30

#### 【0019】

一実施形態において、第1のオフセットは第2のオフセットよりも大きく、第2のオフセットは第3のオフセットよりも大きく、第3のオフセットは第4のオフセットよりも大きい。このようにして、データ処理における後続の各コンポーネントは、ヘッダをオフセットのデータに書き込んで、追加的なヘッダを、それ以前に書き込まれたアプリケーション・データ及びバッファに書き込まれたヘッダ上に配置することができる。

#### 【0020】

他の実施形態において、アプリケーション6は、ストレージ・デバイス16にデータを要求することができる。それに応じて、ストレージ・デバイス・ドライバ14は、当該データにアクセスし、当該データをストレージ・エージェント12に返す。次に、ストレージ・エージェント12(又は他のコンポーネント)は、ストレージ・マネージャAPI8に、検索対象のデータを追加するためのバッファ18を要求することができる。次に、上記の各コンポーネントはオフセット計算を実施して、バッファ18内の場所を設けるオフセットを提供することができ、それによって、各コンポーネントが検索対象のアプリケーション・データに関するヘッダをバッファ18に追加し、その結果、バッファ18を指すポインタが最終的に当該データにアクセスするアプリケーション6に受け渡されるようにすることができる。

40

#### 【0021】

50

様々なコンポーネントは、本明細書に記載の諸実施形態を用いてバッファを共有し、その同一のバッファのデータにアクセスし、ヘッダ情報をその同一のバッファに書き込み、当該バッファを指すポインタを他のコンポーネントに受け渡して、それらの他のコンポーネントが当該バッファにアクセスすることを可能にする。

#### 【0022】

他の実施形態において、アプリケーションは、ストレージ・マネージャ API 8 に複数のバッファを要求し、次にアプリケーション・データ 40 とアプリケーション・ヘッダ 44 のデータを同時に複数のバッファに書き込んだ後、当該バッファのハンドルをストレージ・マネージャ API 8 に受け渡すことができる。

#### 【0023】

##### (追加的な実施形態の詳細)

本明細書に記載の処理は、標準的なプログラミング技法又はエンジニアリング技法あるいはその両方を使用してソフトウェア、ファームウェア、ハードウェア、又はそれらの任意の組合せを製造する方法、装置、又は製品として実施することができる。本明細書で使用する「製品 (article of manufacture)」という用語は、ハードウェア・ロジック（例えば、集積回路チップ、プログラマブル・ゲート・アレイ（PGA）、特定用途向け集積回路（ASIC）等）、あるいは、磁気記憶媒体（例えば、ハード・ディスク・ドライブ、フレキシブル・ディスク、テープ等）、光記憶装置（CD-R OM、光ディスク等）、揮発性及び不揮発性メモリ・デバイス（例えば、EEPROM、ROM、PROM、RAM、DRAM、SRAM、ファームウェア、プログラマブル・ロジック等）のようなコンピュータに読み込み可能な媒体に実装されるコード又はロジックを指す。コンピュータに読み込み可能な媒体内のコードは、プロセッサによってアクセスされ実行される。更に、好ましい諸実施形態がその内部で実装されるコードは、伝送媒体を介してアクセス可能なものとすることができる、あるいはファイル・サーバからネットワークを介してアクセス可能なものとすることができる。そのような場合には、コードがその内部で実装される製品は、ネットワーク伝送回線、無線伝送媒体、空間内を伝搬する信号、電波、赤外線信号等のような伝送媒体を有していてもよい。それ故、「製品」は、コードがその内部で実施される媒体を有していてもよい。更に、「製品」は、コードがその内部で実装され処理され実行されるハードウェア構成要素とソフトウェア構成要素の組合せを有していてもよい。言うまでもなく、当業者なら本発明の範囲から逸脱することなく、多くの修正を上記の構成に施すことができ、上記の製品が当業界で知られる任意の情報保持媒体を有していてもよいことを理解するであろう。

#### 【0024】

図 3 乃至 5 に示される処理は、ある順番で発生する複数のイベントを示している。諸代替実施形態では、上記の処理の一部を別の順番で実施し、修正し、あるいは省略することもできる。また、いくつかのステップが上述のロジックに追加された場合にも、本明細書に記載の諸実施形態と矛盾するものではない。更に、本明細書に記載の処理は逐次的に行うことができ、あるいはいくつかの処理を並列的に処理することができる。更に、各処理は、単一の処理ユニットによって処理されてもよいし、分散処理ユニットによって処理されてもよい。

#### 【0025】

本発明の様々な実施形態の前述の記載は、例示及び説明のために提示してきた。本発明を余すところのないものとし、又は開示の厳密な形態に限定することは、本出願人の意図するところではない。上記の教示に照らせば多くの修正形態及び変更形態が可能である。本発明の範囲は、上記の詳細な説明ではなく添付の特許請求範囲によって限定されることが意図されている。上記の明細書、実施例、及びデータは、本発明の製造及び使用に関する完全な説明を与えるものである。本発明の様々な実施形態は本発明の趣旨及び範囲から逸脱しない範囲で作成することができ、したがって、本発明の範囲は添付の特許請求範囲の各請求項によって定められる。

#### 【図面の簡単な説明】

10

20

30

40

50

## 【0026】

【図1】ネットワーク・コンピューティング環境の一実施形態を示す図である。

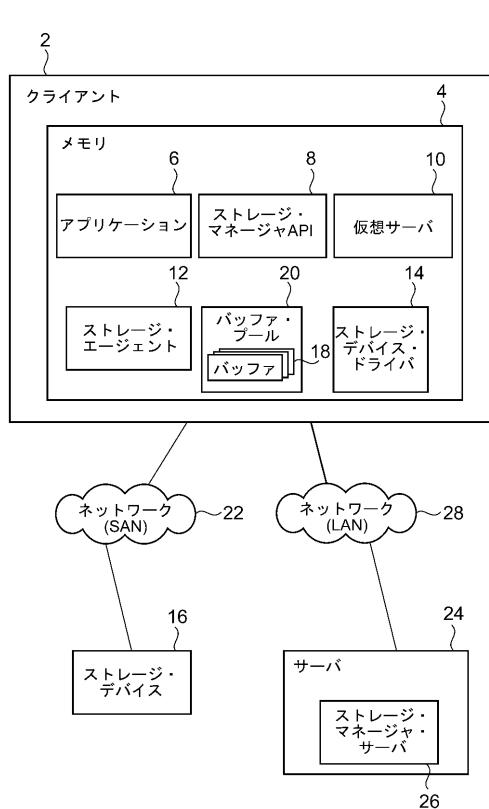
【図2】バッファ・フォーマットの一実施形態を示す図である。

【図3】データをアプリケーションからストレージ・デバイスに転送する処理の一実施形態を示す図である。

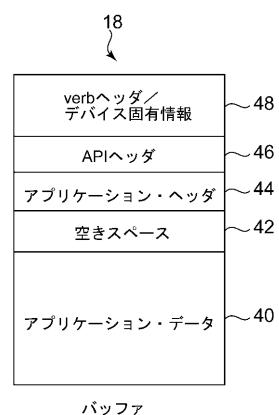
【図4】データをアプリケーションからストレージ・デバイスに転送する処理の一実施形態を示す図である。

【図5】データをアプリケーションからストレージ・デバイスに転送する処理の一実施形態を示す図である。

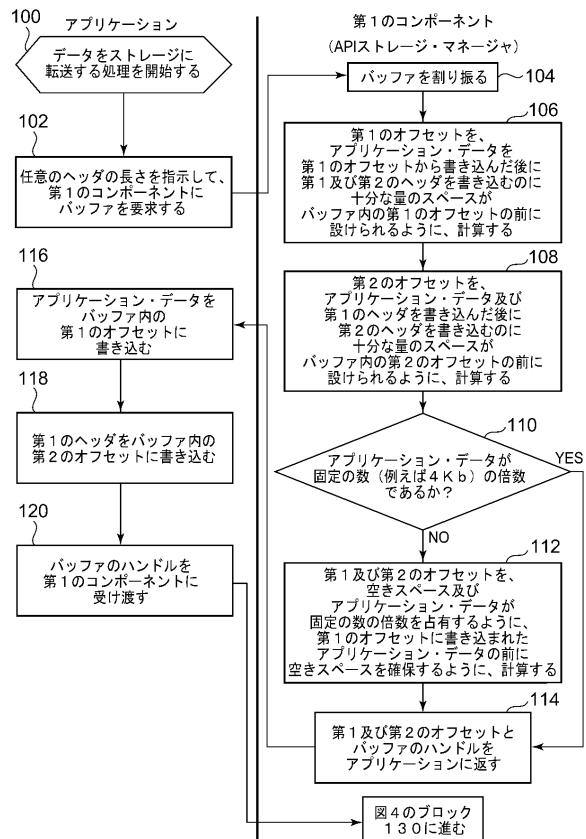
【図1】



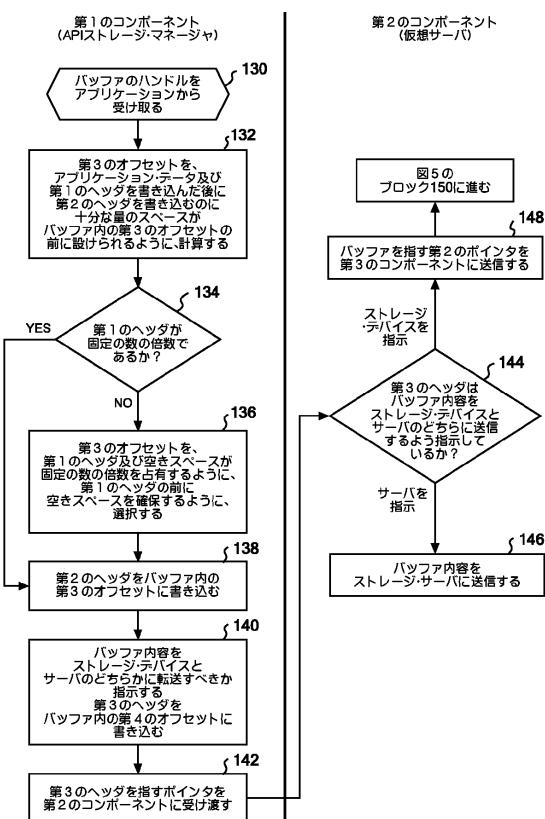
【図2】



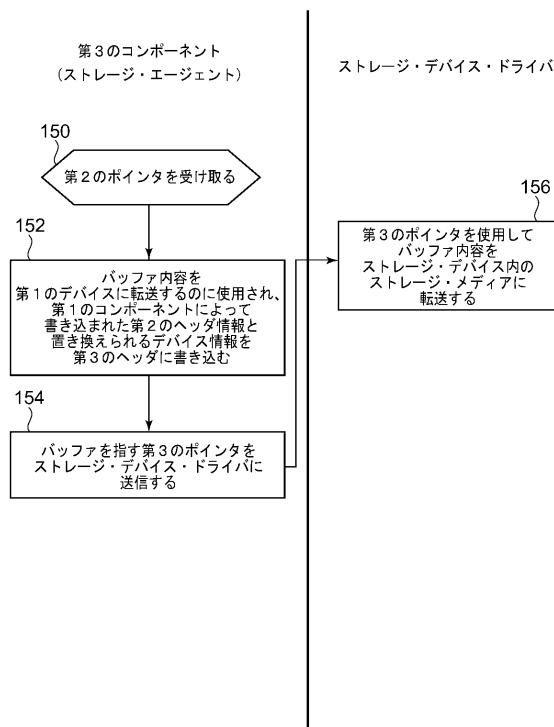
【図3】



【図4】



【図5】



---

フロントページの続き

(74)代理人 100086243

弁理士 坂口 博

(72)発明者 アングリン、マシュー、ジョセフ

アメリカ合衆国85748 アリゾナ州ツーソン ペイジ・クリーク・プレイス1694エス

(72)発明者 ホックバーグ、アビシャイ、ハイム

アメリカ合衆国95123 カリフォルニア州サンノゼ ハックルベリー・ドライブ308

(72)発明者 ヴィクスン、ジョン

アメリカ合衆国94403 カリフォルニア州サン・マテオ メイソン・レーン2672

審査官 池田 聰史

(56)参考文献 特開2004-185349(JP,A)

特開2004-272901(JP,A)

特表平06-505821(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F 12/00

G06F 3/06